شرکت توسعه
فولاد آلیاژی ایران

دانشگاه یزد

سال سوم، شماره دوم،
پاییز و زمستان ۱۴۰۴

پژوهش‌های نظری و کاربردی هوش ماشینی

هوشمندسازی صنعت فولاد: مروری بر کاربردهای عملیاتی هوش مصنوعی در بهینه‌سازی، پایداری و تحول دیجیتال فرآیندهای تولید راضیه شیخ‌پور

پیش‌بینی سری‌های زمانی قیمت فولاد با استفاده از معماری ترنسفورمر پچ‌محور
فاطمه زارع مهرجردی، عباس نوروزی

تشخیص دو پوسته بودن فولاد با شبکه‌های عصبی مصنوعی: رویکردی مبتنی بر بینایی ماشین و یادگیری عمیق
مصطفی مجیدی، سیدامیر مکی نژاد

بهینه‌سازی هوشمند انرژی و کنترل تطبیقی کوره‌های قوس الکتریکی در صنعت فولاد با استفاده از دوقلوی دیجیتال و شبکه‌های عصبی سارا محمودی رشید

طبقه‌بندی عیوب سطح فولاد با مدلی بهبودیافته از معماری VGG16
طوبی ترابی پور ، ابوالفضل گندمی

مدلی ترکیبی در پیش‌بینی داده محور مصرف انرژی در صنعت فولاد: رویکردی بر پایه LSTM و بهینه‌سازی هایپرپارامترها با الگوریتم Optuna لیلا ذوالقدر، مجید شخصی نیائی، یحیی زارع مهرجردی، محمدمهدی لطفی

تشخیص خودکار شمش‌های گرد ناصاف از محصولات خروجی دستگاه صافکاری با استفاده از روش‌های بینایی ماشین راستین نمیرانیان، سحرسلطانی، ولی درهمی

SSLA-Net: معماری یادگیری معنایی ـ مکانی برای شناسایی مقاوم نقص‌های سطحی فولاد
معصومه رضائی، منصوره رضائی، مهری رجائی

هوشمند سازی کنترل کیفیت و تشخیص خرابی در فرآیندهای تولید فولاد با استفاده از یادگیری عمیق و پردازش تصویر به‌منظور بهینه‌سازی مصرف انرژی فاطمه سعادت جو، مسلم ملایی

مدل‌سازی فرآیند در کوره گندله‌سازی با استفاده از روش‌های هوشمند: مطالعه موردی در شرکت معدنی و صنعتی گل‌گهر محدثه رضایی، حسین نظام‌آبادی‌پور، رضا خاکسارپور

تحلیل و شناسایی ویژگی‌های کلیدی برای طبقه‌بندی عیوب سطحی صفحات فولادی با استفاده از مدل‌های یادگیری ماشین و روش متعادل‌سازی داده‌ها حبیب خدادادی، سید علی مهری، مهسا خدادادی

بهبود بخش‌بندی عیوب سطح فولاد با بهینه‌سازی ابرپارامترهای فیلتر گابور توسط الگوریتم ژنتیک و طبقه‌بندی جنگل تصادفی پریسا فتحیان بروجنی، فریبا نمیرانیان، فهیمه رشیدی


بسم الله الرحمن الرحيم

دوفصلنامه علمی
نشریه پژوهش‌های نظری و کاربردی هوش ماشینی

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

شاپا چاپی: ۲۰۲۹-۲۸۲۱
شاپای الکترونیکی: ۲۰۳۷-۲۸۲۱

نشانی: یزد، صفاییه، دانشگاه یزد، پردیس فنی مهندسی، دانشکده مهندسی کامپیوتر، دفتر نشریه پژوهشهای نظری و کاربردی هوش ماشینی. تلفاکس: ۳۱۲۳۳۶۲۸ -۰۳۵۱

پست الکترونیکی: **abmir@journals.yazd.ac.ir**

تارنمای نشریه: **http://abmir.yazd.ac.ir**

فهرست **شماره صفحه**

هوشمندسازی صنعت فولاد: مروری بر کاربردهای عملیاتی هوش مصنوعی در بهینه‌سازی، پایداری و تحول دیجیتال فرآیندهای تولید ۱-۱۶

راضیه شیخ پور

پیش‌بینی سری‌های زمانی قیمت فولاد با استفاده از معماری ترنسفورمر پیچ‌محور ۱۷-۳۳

فاطمه زارع مهرجردی، عباس نوروزی

تشخیص دو پوسته بودن فولاد با شبکه‌های عصبی مصنوعی: رویکردی مبتنی بر بینایی ماشین و یادگیری عمیق ۳۵-۵۲

مصطفی مجیدی، سیدامیر مکی نژاد

بهینه‌سازی هوشمند انرژی و کنترل تطبیقی کوره‌های قوس الکتریکی در صنعت فولاد با استفاده از دوقلوی دیجیتال و شبکه‌های عصبی ۵۳-۷۱

سارا محمودی رشید

طبقه‌بندی عیوب سطح فولاد با مدلی بهبودیافته از معماری VGG16 ۷۳-۸۸

طوبی ترابی پور ، ابوالفضل گندمی

مدلی ترکیبی در پیش‌بینی داده محور مصرف انرژی در صنعت فولاد: رویکردی بر پایه LSTM و بهینه‌سازی هایپرپارامترها با الگوریتم Optuna ۸۹-۱۰۵

لیلا ذوالقدر، مجید شخصی نیائی، یحیی زارع مهرجردی، محمدمهدی لطفی

تشخیص خودکار شمش‌های گرد ناصاف از محصولات خروجی دستگاه صافکاری با استفاده از روش‌های بینایی ماشین ۱۰۷-۱۲۳

راستین نمیرانیان، سحرسلطانی، ولی درهمی

SSLA-Net: معماری یادگیری معنایی ـ مکانی برای شناسایی مقاوم نقص‌های سطحی فولاد ۱۲۵-۱۳۹

معصومه رضائی، منصوره رضائی، مهری رجائی

هوشمند سازی کنترل کیفیت و تشخیص خرابی در فرآیندهای تولید فولاد با استفاده از یادگیری عمیق و پردازش تصویر به‌منظور بهینه‌سازی مصرف انرژی ۱۴۱-۱۵۸

فاطمه سعادت جو، مسلم ملایی

مدل‌سازی فرآیند در کوره گندله‌سازی با استفاده از روش‌های هوشمند: مطالعه موردی در شرکت معدنی و صنعتی گل‌گهر ۱۵۹ - ۱۷۳

محدثه رضایی، حسین نظام‌آبادی‌پور، رضا خاکسارپور

تحلیل و شناسایی ویژگی‌های کلیدی برای طبقه‌بندی عیوب سطحی صفحات فولادی با استفاده از مدل‌های یادگیری ماشین و روش متعادل‌سازی داده‌ها ۱۷۵-۱۸۶

حبیب خدادادی، سید علی مهری، مهسا خدادادی

بهبود بخش‌بندی عیوب سطح فولاد با بهینه‌سازی ابرپارامترهای فیلتر گابور توسط الگوریتم ژنتیک و طبقه‌بندی جنگل تصادفی ۱۸۷-۱۹۷

پریسا فتحیان بروجنی، فریبا نمیرانیان، فهیمه رشیدی

سخن سردبیر

خوشبختانه این شماره نشریه به پژوهش‌های مرتبط با به‌کارگیری هوش مصنوعی در صنعت فولاد اختصاص یافته است. ازاین‌رو مایه خرسندی است که در این شماره، بخش آغازین را به سخنان مدیر عامل محترم شرکت توسعه فولاد آلیاژی ایرانیان اختصاص دهیم:

تحولات شتابان هوش مصنوعی در سال‌های اخیر، چهره صنایع بزرگ جهان را دگرگون کرده است و صنعت فولاد نیز از این جریان تحول‌آفرین بی‌نصیب نمانده است. نیاز روزافزون به افزایش بهره‌وری، ارتقای کیفیت، مدیریت بهینه انرژی و دستیابی به رقابت‌پذیری بیشتر، ما را بر آن می‌دارد که فناوری‌های هوشمند را نه یک انتخاب اختیاری، بلکه ضرورتی راهبردی برای آینده صنعت فولاد کشور بدانیم. به‌درستی گفته شده است که امروز هوش مصنوعی همان نقشی را ایفا می‌کند که برق در سده گذشته ایفا کرد؛ فناوری‌ای که نبود آن، به‌سختی قابل تصور است.

با همین رویکرد، شرکت توسعه فولاد آلیاژی ایرانیان در کنار شرکت فولاد آلیاژی ایران، از انتشار «شماره ویژه کاربردهای سیستم‌های هوشمند در صنعت فولاد» در نشریه پژوهش‌های نظری و کاربردی هوش ماشینی دانشکده مهندسی کامپیوتر دانشگاه یزد حمایت کرده است. هدف این اقدام، ایجاد پیوندی پایدار و کارآمد میان پژوهش‌های دانشگاهی و مسائل واقعی و عملیاتی صنعت فولاد است؛ پیوندی که بتواند دستاوردهای علمی را به راه‌حل‌هایی اجرایی و اثربخش برای کارخانه‌های فولادی تبدیل کند.

در این همکاری، محورهای اصلی فراخوان مقالات بر اساس نیازهای واقعی واحدهای تولیدی و ظرفیت‌های هوش مصنوعی در حل چالش‌های عملیاتی تعیین شد تا این شماره ویژه صرفاً جنبه نظری نداشته باشد و نتایج آن بتواند مستقیماً در محیط صنعت مورد استفاده قرار گیرد. فرآیند داوری نیز با مشارکت تیمی مشترک از اعضای هیأت علمی دانشگاه‌ها و متخصصان صنعت فولاد و مطابق استانداردهای علمی نشریه انجام شد و در نهایت دوازده مقاله برای انتشار برگزیده شد. این مجموعه، نمونه‌ای ارزشمند از هم‌افزایی میان دانشگاه و صنعت و نشانه‌ای از ظرفیت بالای کشور در توسعه کاربردهای هوش مصنوعی در صنعت فولاد است.

انتشار این شماره ویژه را گامی مؤثر در مسیر تحول دیجیتال صنعت فولاد می‌دانیم. امید است نتایج حاصل بتواند در ارتقای بهره‌وری، کاهش هزینه‌ها، بهبود کیفیت و حل مسائل پیچیده خطوط تولید نقشی ملموس ایفا کند. همچنین انتظار می‌رود این همکاری سازنده میان صنعت و دانشگاه در سال‌های آینده استمرار یابد و زمینه‌ساز توسعه پژوهش‌های کاربردی و تربیت متخصصانی شود که بتوانند در مسیر هوشمندسازی صنعت فولاد نقش‌آفرینی کنند. در پایان، از همکاران نشریه در دانشگاه یزد، همکاران در شرکت توسعه فولاد آلیاژی ایرانیان، داوران ارجمند، متخصصان صنعت فولاد و همه پژوهشگرانی که در این فراخوان مشارکت کردند، صمیمانه قدردانی می‌کنم. اعتماد و تلاش مشترک ما سرمایه‌ای ارزشمند برای آینده صنعت فولاد کشور است.

ولی درهمی
سردبیر نشریه

احسان حسینی
مدیرعامل شرکت توسعه فولاد آلیاژی ایرانیان

شرکت توسعه فولاد آلیاژی ایرانیان

دانشگاه یزد

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱٤۰٤، صفحات ۱٦-۱

هوشمندسازی صنعت فولاد: مروری بر کاربردهای عملیاتی هوش مصنوعی در بهینه‌سازی، پایداری و تحول دیجیتال فرآیندهای تولید

راضیه شیخ‌پور*

دانشیار، گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه اردکان، اردکان، ایران

مقاله پژوهشی

تاریخ دریافت:
۱٤۰٤/۰٤/۱۵

تاریخ پذیرش:
۱٤۰٤/۰۷/۲۰

کلیدواژه‌ها:
هوش مصنوعی، صنعت فولاد، یادگیری ماشین، بهینه‌سازی فرآیند، تحول دیجیتال

نویسنده‌مسئول:
rsheikhpour@ardakan.ac.ir

چکیده

صنعت فولاد به‌عنوان یکی از ارکان حیاتی اقتصاد جهانی، با چالش‌هایی نظیر نوسانات بازار، افزایش هزینه‌های انرژی، ضرورت کاهش آلاینده‌ها و رقابت فزاینده در بهره‌وری مواجه است. در این زمینه، هوش مصنوعی به‌عنوان محرکی کلیدی در تحول دیجیتال، نقش مهمی در بهینه‌سازی و هوشمندسازی فرآیندهای تولید ایفا می‌کند. این مقاله مروری، به بررسی کاربردهای عملیاتی هوش مصنوعی در صنعت فولاد می‌پردازد. در ابتدا، مفاهیم پایه‌ای هوشمندسازی صنعتی و الگوریتم‌های هوش مصنوعی معرفی شده‌اند. سپس، کاربردهای متنوع این فناوری از جمله بهینه‌سازی مصرف انرژی، کنترل کیفیت، نگهداری پیش‌بینانه، پیش‌بینی خواص مواد، تصمیم‌گیری در زنجیره تأمین و اتوماسیون خطوط تولید مورد بررسی قرار گرفته و مزایای پیاده‌سازی هوش مصنوعی بیان شده است. در ادامه، چالش‌های پیاده‌سازی این فناوری از جمله نیاز به داده‌های دقیق، زیرساخت‌های فناورانه، ملاحظات امنیتی و مسئله تفسیرپذیری مدل‌ها تحلیل شده است. در پایان نیز، روندهای نوظهوری نظیر ادغام هوش مصنوعی با فناوری‌هایی چون دوقلوی دیجیتال، یادگیری فدرال و مدل‌های قابل تفسیر معرفی شده‌اند. نتایج بررسی‌ها نشان می‌دهد بهره‌گیری هدفمند از هوش مصنوعی نه‌تنها به افزایش بهره‌وری و کاهش هزینه‌ها کمک می‌کند، بلکه مسیر دستیابی به تولید پایدار و واکنش‌پذیر را نیز هموار می‌سازد.

: 10.22034/ABMIR.2025.23461.1146

E-ISSN: 2821-2037

۱- مقدمه

صنعت فولاد، به‌عنوان ستون فقرات توسعه صنعتی، نقشی راهبردی در رشد اقتصادی کشورها ایفا می‌کند. این صنعت با تأمین مواد اولیه کلیدی برای بخش‌هایی چون ساختمان‌سازی، خودروسازی، انرژی، حمل‌ونقل، پتروشیمی و حتی فناوری‌های نوین، جایگاهی حیاتی در زنجیره ارزش جهانی دارد [۱–۳]. با این حال، فرآیندهای پیچیده، پیوسته و سنگین در تولید فولاد، همواره با چالش‌هایی جدی از جمله مصرف بالای انرژی، نوسانات کیفیت، وابستگی شدید به تجهیزات سنتی و آسیب‌پذیری بالا در برابر توقف‌های ناگهانی مواجه بوده‌اند. بهره‌برداری از منابع انرژی پرهزینه مانند برق، زغال‌سنگ و گاز طبیعی، ضمن افزایش هزینه‌های تولید، اثرات زیست‌محیطی قابل‌توجهی نیز به‌دنبال داشته است. همچنین، نبود سیستم‌های پایش هوشمند و تصمیم‌گیری بلادرنگ، کنترل کیفیت و نگهداری تجهیزات را با چالش‌های عملیاتی مواجه کرده است [۲–٤].

در پاسخ به این چالش‌ها، حرکت به‌سوی هوشمندسازی فرآیندهای تولید فولاد به ضرورتی استراتژیک تبدیل شده است. با ظهور انقلاب صنعتی چهارم و رشد فناوری‌هایی نظیر اینترنت اشیاء صنعتی، تحلیل داده‌های بزرگ، بینایی ماشین و به‌ویژه هوش مصنوعی، کارخانه‌های فولاد به سمت «فولادسازی هوشمند» گام برداشته‌اند. در این کارخانه‌ها، داده‌ها از طریق حسگرهای پیشرفته گردآوری و توسط الگوریتم‌های هوش مصنوعی تحلیل می‌شوند تا الگوها، ناهنجاری‌ها و فرصت‌های بهینه‌سازی شناسایی گردند [٤–٦]. چارچوبی چهارلایه برای پیاده‌سازی فناوری تولید هوشمند، شامل لایه تجهیزات، شبکه، داده‌های حجیم و لایه کاربردی، زیربنای این تحول فناورانه را شکل می‌دهد [٦].

هوش مصنوعی، به‌عنوان یکی از فناوری‌های کلیدی در این مسیر، توانسته است با بهره‌گیری از قدرت یادگیری و تحلیل بر پایه داده، ابزارهایی مؤثر برای پیش‌بینی، بهینه‌سازی و تصمیم‌سازی ارائه دهد. کاربردهای آن در صنعت فولاد طیف وسیعی را دربر می‌گیرد: از پیش‌بینی شرایط بهینه تولید و کنترل کیفیت پیشرفته با بینایی ماشین گرفته تا تعمیرات پیش‌بینانه و کاهش ردپای زیست‌محیطی از طریق بهینه‌سازی مصرف انرژی. این فناوری از یک ابزار فناورانه صرف، به یک مزیت رقابتی و الزام استراتژیک در صنعت فولاد جهانی تبدیل شده است [۵–۸].

در این مقاله، با هدف ترسیم چشم‌اندازی روشن و کاربردی، یک مروری بر کاربردهای عملیاتی هوش مصنوعی در صنعت فولاد ارائه شده است. تمرکز اصلی بر نقش این فناوری در بهینه‌سازی بهره‌وری، ارتقای کیفیت، کاهش هزینه‌ها و پشتیبانی از تصمیم‌گیری‌های بلادرنگ در فرآیندهای پیچیده و حساس تولید است. مطالعه حاضر ضمن بررسی روش‌ها و الگوریتم‌های هوش مصنوعی در حوزه‌های مختلف زنجیره ارزش فولاد، مزایا و دستاوردهای کاربرد هوش مصنوعی را تحلیل می‌کند. همچنین، چالش‌ها و موانع فنی، داده‌ای، زیرساختی و امنیتی مرتبط با پیاده‌سازی این فناوری‌ها نیز مورد بررسی قرار گرفته است. در بخش پایانی، روندهای نوظهور و فناوری‌های مکمل معرفی شده و چشم‌اندازهای آینده هوشمندسازی پایدار و واکنش‌پذیر در صنعت فولاد ترسیم می‌شود.

این مقاله با هدف پاسخ به پرسش‌های پژوهشی زیر طراحی شده است:

۱- مهم‌ترین حوزه‌های کاربردی هوش مصنوعی در صنعت فولاد کدام‌اند و چه الگوریتم‌هایی بیشترین استفاده را دارند؟
۲- هوش مصنوعی چگونه به بهینه‌سازی فرآیندها، افزایش بهره‌وری و ارتقای کیفیت محصولات فولادی کمک کرده است؟
۳- چه چالش‌ها و موانعی در مسیر پیاده‌سازی راهکارهای مبتنی بر هوش مصنوعی در صنعت فولاد وجود دارد؟
٤- روندهای نوظهور و مسیرهای آینده‌پژوهی در این حوزه کدام‌اند و چگونه می‌توانند خلأهای موجود در ادبیات را پوشش دهند؟

این مقاله با تمرکز بر مطالعات اخیر، تلاش می‌کند تصویری جامع از وضعیت کنونی و چشم‌انداز آینده هوش مصنوعی در صنعت فولاد ارائه دهد. بدین ترتیب، خلأ ناشی از فقدان یک چارچوب یکپارچه برای تحلیل کاربردهای عملیاتی، موانع اجرایی و روندهای آینده این فناوری در صنعت فولاد برطرف می‌شود.

ساختار ادامه مقاله به شرح زیر سازماندهی شده است: در بخش ۲، مبانی و مفاهیم هوش مصنوعی تشریح می‌شود. بخش ۳ به بررسی کاربردهای عملیاتی هوش مصنوعی در فرآیندهای تولید فولاد می‌پردازد. در بخش ٤، مزایا و دستاوردهای استفاده از هوش مصنوعی در این صنعت مورد تحلیل قرار می‌گیرد. بخش ۵ به چالش‌ها و موانع پیاده‌سازی فناوری‌های هوش مصنوعی در صنعت فولاد می‌پردازد و در بخش ٦، روندهای نوظهور و چشم‌اندازهای آینده هوش مصنوعی در این حوزه بررسی می‌شوند. نهایتاً، در بخش ۷، نتایج کلی پژوهش ارائه می‌شود.

۲- مبانی هوشمندسازی و هوش مصنوعی

در این بخش، مبانی هوشمندسازی صنعتی و الگوریتم‌های هوش مصنوعی معرفی می‌شوند تا زمینه لازم برای بررسی کاربردهای آن‌ها در صنعت فولاد فراهم شود.

۲-۱ تعریف هوشمندسازی صنعتی

هوشمندسازی صنعتی به معنای ادغام فناوری‌های نوین دیجیتال با فرآیندهای تولید، نگهداری و مدیریت زنجیره تأمین با هدف ارتقای بهره‌وری، دقت، انعطاف‌پذیری و سرعت در تصمیم‌گیری‌های صنعتی است. در این چارچوب، تجهیزات سنتی با بهره‌گیری از حسگرها، ارتباطات اینترنتی و الگوریتم‌های پیشرفته، به سامانه‌هایی هوشمند، خودآموز و پاسخ‌گو تبدیل می‌شوند که قادرند به‌صورت بلادرنگ داده‌ها را جمع‌آوری کرده، تحلیل نموده و تصمیم‌سازی بهینه انجام دهند [٤، ٦، ۹].

به‌طور کلی، پیاده‌سازی تولید هوشمند نیازمند معماری‌ای چندلایه است که مراحل گردآوری تا بهره‌برداری از داده را در بردارد [٦]. شکل ۱ این معماری را در قالب چهار لایه اصلی نمایش می‌دهد.

۲-۲ الگوریتم‌های هوش مصنوعی در تولید صنعتی

در صنعت فولاد، به‌دلیل پیچیدگی فرآیندها، تغییرپذیری شرایط عملیاتی و نیاز به بهینه‌سازی مستمر، استفاده از الگوریتم‌های هوش مصنوعی اهمیت ویژه‌ای دارد. این الگوریتم‌ها با تحلیل داده‌های متنوع و بزرگ، به بهبود کیفیت محصولات، افزایش کارایی تولید، پیش‌بینی خطاها و کاهش هزینه‌ها کمک می‌کنند [٤]. در ادامه، مهم‌ترین الگوریتم‌های هوش مصنوعی و کاربردهای آن‌ها در صنعت فولاد معرفی می‌شوند.

شکل (۱): معماری پیاده‌سازی تولید هوشمند در صنعت فولاد

۲-۲-۱ یادگیری ماشین

یادگیری ماشین شامل الگوریتم‌هایی است که بدون نیاز به برنامه‌نویسی صریح، از داده‌ها الگوها را استخراج و تصمیم‌گیری انجام می‌دهند. این الگوریتم‌ها برای مدل‌سازی روابط پیچیده و غیرخطی میان متغیرهای تولیدی بسیار مناسب هستند.

کاربردها در صنعت فولاد [۱۰، ۱۱]:

- پیش‌بینی پارامترهای مهم مانند دما، فشار و نرخ تولید در کوره‌ها
- طبقه‌بندی محصولات بر اساس ترکیب شیمیایی و خواص مکانیکی
- تحلیل مصرف انرژی و شناسایی نقاط اتلاف

الگوریتم‌های رایج: جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM)، k-نزدیک‌ترین همسایه (KNN)، رگرسیون خطی (LR). دلیل استفاده گسترده از الگوریتم‌های یادگیری ماشین در صنعت فولاد، توانایی آن‌ها در مدل‌سازی روابط پیچیده و غیرخطی

میان متغیرهای فرآیند و همچنین مدیریت داده‌های نویزی و متنوع است.

مثال عملی: در یک مطالعه صنعتی در شرکت فولاد خوزستان، از رگرسیون خطی چندگانه، رگرسیون غیرخطی چندگانه و شبکه عصبی پرسپترون برای توسعه مدلی هوشمند به‌منظور پیش‌بینی شاخص خوردگی و رسوب در مدارهای آب خنک‌کننده کوره‌های قوس الکتریکی استفاده شد [۱۲].

۲-۲-۲ شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی ساختاری الهام‌گرفته از مغز انسان دارند و قادر به یادگیری روابط پیچیده و غیرخطی بین داده‌ها هستند. این شبکه‌ها برای مدل‌سازی فرآیندهای پیچیده صنعتی بسیار کارآمد هستند.

کاربردها در صنعت فولاد [۱۳]:

- مدل‌سازی فرآیندهای غیرخطی نظیر نورد گرم و سرد، عملیات حرارتی و ریخته‌گری پیوسته
- کنترل کیفیت محصولات از طریق یادگیری روابط ورودی-خروجی
- تشخیص نواقص داخلی و سطحی فولاد با استفاده از داده‌های تصویری یا حسگری

الگوریتم‌های رایج: پرسپترون چند لایه ((MLP، توابع شعاعی پایه (RBF). این شبکه‌ها به‌دلیل قابلیت بالای خود در یادگیری روابط غیرخطی و چندمتغیره، برای مدل‌سازی فرآیندهای پیچیده‌ای مانند نورد و عملیات حرارتی در فولادسازی بسیار مناسب هستند.

مثال عملی: در یک مطالعه، از شبکه عصبی پیش‌خور برای توسعه مدلی به‌منظور پیش‌بینی ضخامت ورق‌های نورد ترکیبی تیتانیوم/فولاد استفاده شد. این مدل با بهره‌گیری از داده‌های ترکیب شیمیایی فولاد و پارامترهای فرآیند نورد، توانست ضخامت نهایی ورق‌ها را با دقت بالایی پیش‌بینی کند. این روش در محیط آزمایشگاهی فرآیند نورد گرم مورد استفاده قرار گرفت و نشان داد که شبکه‌های عصبی قابلیت بالایی در مدل‌سازی روابط پیچیده بین متغیرهای فرآیندی دارند [۱۴].

۲-۲-۳ یادگیری عمیق

یادگیری عمیق شاخه‌ای از شبکه‌های عصبی است که با استفاده از معماری‌های چندلایه، قابلیت بالایی در تحلیل داده‌های پیچیده و حجیم دارد. این رویکرد به‌ویژه در پردازش تصاویر، صوت و داده‌های سری زمانی عملکرد برجسته‌ای نشان داده است. هرچند یادگیری عمیق به‌طور کلی زیرمجموعه‌ای از شبکه‌های عصبی محسوب می‌شود، اما به‌دلیل تفاوت‌های اساسی در معماری، الگوریتم‌ها و همچنین حجم قابل‌توجه پژوهش‌های اخیر در صنعت فولاد، در این مقاله به‌عنوان بخشی مستقل از شبکه‌های عصبی کلاسیک مانند MLP و RBF بررسی شده است.

کاربردها در صنعت فولاد [۵، ۱۵، ۱۶]:

- شناسایی خودکار عیوب سطحی مانند ترک، خراش و موج‌دار بودن
- پیش‌بینی رفتار حرارتی مواد در کوره‌ها
- مدل‌سازی خواص مکانیکی آلیاژهای جدید

الگوریتم‌های رایج: شبکه‌های عصبی کانولوشن (CNN) برای تصاویر، شبکه‌های عصبی بازگشتی (RNN)، شبکه‌های حافظه طولانی کوتاه مدت (LSTM) برای داده‌های سری زمانی. الگوریتم‌های یادگیری عمیق به‌ویژه برای داده‌های حجیم تصویری و زمانی در صنعت فولاد انتخاب می‌شوند، زیرا در شناسایی عیوب سطحی و تحلیل رفتار دینامیک فرآیندها دقت بالایی دارند.

مثال عملی: در یک مطالعه آزمایشگاهی، از شبکه CNN بر روی مجموعه داده NEU-CLS برای تشخیص عیوب سطحی فولاد شامل ترک‌ها و ناپیوستگی‌های سطح استفاده شد [۱۵]. همچنین شرکت POSCO کره جنوبی، مدل‌های یادگیری عمیق را در فرآیند پوشش‌دهی ورق‌های فولادی به‌کار گرفته است تا کنترل دقیقی بر پارامترهایی مانند ضخامت فولاد، عرض، شرایط عملیاتی و میزان پوشش هدف فراهم شود [۱۶].

یادگیری عمیق در سال‌های اخیر به‌طور مؤثر در بهینه‌سازی فرآیندها، شناسایی عیوب و پیش‌بینی خواص مواد به‌کار رفته است. جدول ۱ کاربردهای کلیدی این روش‌ها را نشان می‌دهد.

جدول (۱): مقایسه روش‌های یادگیری عمیق در کاربردهای کلیدی در صنعت فولاد

کاربرد	نوع داده	مدل اصلی	هدف	مزایا یا نتیجه	محل اجرا/ داده‌ها	مرجع
پیش‌بینی کیفیت کک	داده برداری	MLP	پیش‌بینی CSR	دقت بالا	داده‌های یک کارخانه کک‌سازی در استرالیا	[۱۷]
پیش‌بینی FeO در سینتر	داده زمانی	LSTM	ترکیب بهینه	دقت زیاد، کاهش خطا	داده‌های کارخانه سینتر (چین)	[۱۸، ۱۹]
تشخیص شکست در ریخته‌گری	دما / زمان	شبکه عصبی BP	تشخیص شکست	کاهش false alarm	داده‌های یک کارخانه فولاد در چین	[۲۰، ۲۱]
تشخیص ترک سطحی	تصویر NEU	YOLO, R-CNN	طبقه‌بندی عیوب	دقت بالا	مجموعه‌داده استاندارد NEU-DET	[۲۲، ۲۳]
شناسایی ساختارهای فولاد	OM / SEM	U-net CNN	طبقه‌بندی/ سگمنتیشن	تطابق با تصاویر آزمایشگاهی	داده‌های آزمایشگاهی (کره جنوبی)	[۲۴، ۲۵]
پیش‌بینی دمای BOF مذاب	طیف شعله	MLP	کنترل نقطه پایان	آموزش‌پذیری سریع	داده‌های یک کارخانه فولاد در چین	[۱۳]
تسریع روش فاز فیلد	داده شبیه‌سازی	PCA + LSTM PCA+GRU	کاهش هزینه محاسباتی	افزایش سرعت تا ۵۰،۰۰۰ برابر	مطالعه شبیه‌سازی (آمریکا)	[۲۶]
تولید داده مصنوعی برای آموزش مدل‌ها	تصویر NEU	GAN	تقویت داده آموزشی	بهبود تشخیص عیوب	تولید داده برای مجموعه NEU(چین)	[۲۷]

۲-۲-٤ بینایی ماشین

بینایی ماشین با تحلیل خودکار تصاویر صنعتی، امکان بازرسی دقیق کیفیت محصولات را در خطوط تولید فولاد فراهم می‌کند.

کاربردها در صنعت فولاد [۱، ۲۸]:

- شناسایی عیوب سطحی در خطوط نورد و ریخته‌گری
- کنترل دقیق ابعاد قطعات تولیدی
- نظارت ایمنی بر کارکنان از طریق تحلیل ویدئو

فناوری‌های رایج: شناسایی اشیا، بخش‌بندی تصویر. علت اصلی انتخاب فناوری‌های بینایی ماشین در صنعت فولاد، نیاز به بازرسی سریع و دقیق کیفیت سطح محصولات در خطوط تولید است.

مثال عملی: در یک مطالعه صنعتی، از دوربین‌های صنعتی و معماری بهبودیافته YOLOv5 برای شناسایی عیوب زنگ‌زدگی در سطح ورق‌های فولادی گالوانیزه استفاده شد [۲۸]. همچنین، شرکت POSCO کره جنوبی سامانه بینایی ماشین را در خطوط تولید خود برای بازرسی سطح ورق‌های فولادی و طبقه‌بندی عیوب سطحی به‌کار گرفته است [۱٦].

۲-۲-۵ الگوریتم‌های خوشه‌بندی و طبقه‌بندی

این الگوریتم‌ها با تحلیل داده‌های حجیم به استخراج الگوها کمک می‌کنند؛ طبقه‌بندی برای داده‌های برچسب‌دار و خوشه‌بندی برای داده‌های بدون برچسب به‌کار می‌رود.

کاربردها در صنعت فولاد [۲۹، ۳۰]:

- دسته‌بندی محصولات نهایی بر اساس ترکیب شیمیایی، ابعاد یا خواص مکانیکی
- جداسازی داده‌های نرمال از داده‌های نویزدار در سیستم‌های جمع‌آوری داده
- تشخیص رفتارهای غیرعادی تجهیزات برای هشدار زودهنگام

الگوریتم‌های رایج: خوشه‌بندی سلسله‌مراتبی، K-means و درخت‌های تصمیم. این الگوریتم‌ها به دلیل توانایی در دسته‌بندی محصولات متنوع فولادی و تشخیص داده‌های غیرعادی، انتخاب مناسبی برای مسائل کیفیت و کنترل فرآیند هستند.

مثال عملی: در یک مطالعه آزمایشگاهی، با استفاده از الگوریتم K-medoids و شبکه LSTM، متغیرهای فرآیند فولادسازی پیش‌بینی

و خوشه‌بندی شدند تا داده‌های واقعی فرآیند فولادسازی تحلیل و دسته‌بندی گردند [۳۰].

۲-۲-۶ الگوریتم‌های بهینه‌سازی

الگوریتم‌های بهینه‌سازی به صورت خودکار پارامترهای بهینه برای کاهش هزینه، مصرف انرژی و افزایش بازدهی را تعیین می‌کنند. کاربردها در صنعت فولاد [۳۱، ۳۲]:

- تنظیم دقیق پارامترهای کوره‌ها، از جمله دما و سرعت شارژ
- زمان‌بندی بهینه تعمیرات و توقفات برنامه‌ریزی شده
- بهینه‌سازی مصرف انرژی در پمپ‌ها و خطوط انتقال

الگوریتم‌های رایج: الگوریتم ژنتیک (GA)، بهینه‌سازی ازدحام ذرات (PSO)، بهینه‌سازی کلونی مورچه‌ها. الگوریتم‌های بهینه‌سازی به دلیل قابلیت حل مسائل چندهدفه و پیچیده، در بهینه‌سازی مصرف انرژی و زمان‌بندی تولید در فولادسازی بسیار پرکاربرد هستند.

مثال عملی: در یک مطالعه آزمایشگاهی، از الگوریتم ترکیبی ژنتیک و بهینه‌سازی اجتماع ذرات (GA-PSO) برای کنترل پروفیل ضخامت عرضی ورق‌های فولادی الکتریکی در فرآیند نورد سرد ۶-استاند استفاده شد تا یکنواختی ضخامت ورق‌ها بهبود یابد و خطای تولید کاهش پیدا کند [۳۲].

۳- کاربردهای عملیاتی هوش مصنوعی در صنعت فولاد

صنعت فولاد، به‌واسطه ماهیت پیچیده، تغییرپذیری پویا و وابستگی شدید به داده‌های فرآیندی، بستری بسیار مناسب برای بهره‌گیری از فناوری‌های هوش مصنوعی فراهم کرده است. استفاده عملیاتی از هوش مصنوعی در این صنعت، با هدف ارتقای بهره‌وری، کاهش هزینه‌ها، بهبود کیفیت محصولات و دستیابی به پایداری زیست‌محیطی، به‌طور فزاینده‌ای در حال گسترش است [۴- ۷].

بررسی‌های اخیر نشان می‌دهند که به‌کارگیری الگوریتم‌های یادگیری ماشین و یادگیری عمیق در فرآیندهای مختلف زنجیره فولاد، نتایج چشمگیری به همراه داشته است [۳، ۷]. برای نمونه، مطالعه‌ای [۵] به بررسی نقش فناوری‌های نوین، به‌ویژه هوش مصنوعی، یادگیری ماشین و یادگیری عمیق در ارتقای بهره‌وری و پایداری زنجیره ارزش فولاد پرداخته است. این مطالعه، در چارچوب مفاهیم صنعت ۴,۰ و تحول دیجیتال، تمرکز خود را بر سامانه‌هایی همچون پایش هوشمند، کنترل پیشرفته، نگهداری پیش‌بینانه و سیستم‌های تصمیم‌یار زیست‌محیطی قرار داده است؛ سامانه‌هایی که با تحلیل داده‌های عملیاتی نظیر دما، فشار، تصاویر حرارتی و داده‌های کیفی محصولات، امکان تصمیم‌گیری هوشمند و دقیق در سطوح مختلف عملیاتی را فراهم می‌کنند. به‌کارگیری روش‌هایی نظیر درخت تصمیم، شبکه‌های عصبی کانولوشنی و یادگیری تقویتی (RL) در زمینه‌هایی همچون بهینه‌سازی مصرف انرژی، کاهش انتشار گازهای گلخانه‌ای به‌ویژه دی‌اکسیدکربن (CO_2) و افزایش کارایی تجهیزات کلیدی نظیر کوره‌های بازگرم، موجب ارتقای دقت پیش‌بینی، کاهش خطاها و بهبود چشمگیر در هزینه‌های عملیاتی شده است. نتایج این مطالعات همچنین حاکی از برتری قابل توجه مدل‌های یادگیری عمیق در مقایسه با روش‌های سنتی، به‌ویژه در تحلیل داده‌های پیچیده و با ابعاد بالا است.

در ادامه، مهم‌ترین کاربردهای عملیاتی هوش مصنوعی در مراحل مختلف زنجیره تولید فولاد به‌صورت طبقه‌بندی‌شده مرور می‌شود.

۳-۱ بهینه‌سازی مصرف انرژی

صنعت فولاد از جمله صنایع انرژی‌بر به‌شمار می‌رود، به‌گونه‌ای که بخش عمده‌ای از هزینه‌های عملیاتی آن به مصرف انرژی، به‌ویژه در تجهیزاتی مانند کوره‌های پیش‌گرم، قوس الکتریکی و خطوط نورد اختصاص دارد. در این زمینه، هوش مصنوعی با فراهم‌سازی ابزارهای پیش‌بینی دقیق و تصمیم‌گیری مبتنی بر داده، امکان بهینه‌سازی مصرف انرژی را فراهم کرده است [۸، ۳۳]. برای نمونه، در مطالعه‌ای [۳۴]، با هدف پیش‌بینی مصرف انرژی، عملکرد چندین الگوریتم یادگیری شامل شبکه عصبی مصنوعی، نزدیک‌ترین همسایه، جنگل تصادفی، گرادیان افزایشی و یک مدل یادگیری عمیق مبتنی بر حافظه بلندمدت (LSTM) مقایسه شد. نتایج نشان داد مدل LSTM با خطای پیش‌بینی پایین‌تر عملکرد برتری دارد. همچنین، الگوریتم‌های گروهی مانند جنگل تصادفی و گرادیان افزایشی نیز دقت مناسبی ارائه دادند، در حالی‌که

مدل‌های ساده‌تر از توانایی کمتری در مدل‌سازی رفتار انرژی برخوردار بودند.

۳-۲ کنترل کیفیت و تشخیص عیوب

کیفیت سطحی و ساختار داخلی محصولات فولادی، یکی از عوامل کلیدی در پذیرش بازار و عملکرد محصول نهایی به‌شمار می‌رود. نوسانات در پارامترهای فرآیند می‌توانند منجر به بروز عیوبی مانند ترک، حفره یا ناخالصی شوند که در صورت عدم شناسایی به‌موقع، موجب کاهش کیفیت یا رد شدن محصول می‌گردند. در این راستا، کاربرد الگوریتم‌های یادگیری عمیق و بینایی ماشین تحولی چشمگیر در سامانه‌های بازرسی کیفیت ایجاد کرده است. این سامانه‌ها قادرند به‌صورت بلادرنگ تصاویر با وضوح بالا از نوارهای فولادی را تحلیل کرده و عیوب سطحی مختلفی همچون ترک، لکه، ناهمواری و موج‌دار بودن را با دقت بالا شناسایی کنند [۲، ۲۳، ۲۷]. به‌عنوان نمونه، در مطالعه‌ای [۱] با استفاده از روش ELA-YOLO مبتنی بر معماری بهینه‌شده YOLOv8، سیستمی دقیق و سریع برای تشخیص عیوب سطحی توسعه داده شد که عملکردی بهینه در مقایسه با مدل‌های پیشین نشان داده است. آزمایش‌های انجام‌شده بر روی مجموعه‌داده‌های صنعتی استاندارد کارایی بالای این روش را در تشخیص عیوب کوچک و پیچیده تأیید می‌کند.

۳-۳ بهینه‌سازی فرآیندهای تولید

فرآیندهای تولید فولاد به‌دلیل پیچیدگی بالا و ماهیت پویا، نیازمند سامانه‌های کنترلی هوشمند برای تنظیم دقیق پارامترهای عملیاتی هستند. هوش مصنوعی، به‌ویژه الگوریتم‌های یادگیری ماشین و یادگیری تقویتی، در بهینه‌سازی خودکار پارامترهایی چون دمای کوره، نرخ تزریق اکسیژن و سرعت غلتک‌های نورد نقش برجسته‌ای ایفا می‌کند. برای نمونه، شرکت Tata Steel در هند با بهره‌گیری از مدل‌های هوش مصنوعی، توانسته خطوط تولید را با تحلیل داده‌های گذشته و بلادرنگ بهینه نماید [۳۵]. در مطالعه‌ای دیگر [۱۱]، با هدف بهینه‌سازی عملکرد کوره قوس الکتریکی، مدلی مبتنی بر الگوریتم رگرسیون بردار پشتیبان توسعه یافته است که دمای نهایی مذاب را در زمان واقعی پیش‌بینی کرده و توان ورودی را به‌صورت خودکار تنظیم می‌کند. نتایج این مدل کاهش ۱۷٪ در انحراف دما و صرفه‌جویی ۲۸۲ کیلووات‌ساعت انرژی در هر مرحله ذوب را نشان می‌دهد. همچنین، نرخ بازگشت سرمایه‌گذاری ۳۵/۸٪ گزارش شده که نشانگر مزیت اقتصادی و زیست‌محیطی این راهکار است.

۳-۴ پیش‌بینی عملکرد و نگهداری پیش‌بینانه تجهیزات

نگهداری پیش‌بینانه از کاربردهای کلیدی هوش مصنوعی در صنعت فولاد است که با هدف پیشگیری از خرابی‌های ناگهانی و کاهش زمان توقف تولید به‌کار می‌رود. در این رویکرد، داده‌های جمع‌آوری‌شده از حسگرهای نصب‌شده روی تجهیزات (مانند یاتاقان‌ها، موتورها و پمپ‌ها) شامل ارتعاش، دما، صدا و فشار، توسط الگوریتم‌های یادگیری ماشین تحلیل می‌شوند تا خرابی‌های بالقوه در مراحل اولیه شناسایی شوند [۵، ۳۶]. در یک مطالعه مرور سیستماتیک [۴]، ۲۱۹ تحقیق در زمینه استفاده از یادگیری ماشین و یادگیری عمیق در نگهداری تجهیزات خطوط فولادسازی (به‌ویژه کوره بلند و نورد گرم) بررسی شده‌اند. داده‌های حسگرها برای تشخیص خرابی و پیش‌بینی عمر باقی‌مانده تجهیزات به‌کار رفته و مدل‌هایی مانند ماشین بردار پشتیبان، جنگل تصادفی و شبکه‌های عصبی پیشرفته نظیر LSTM مورد ارزیابی قرار گرفته‌اند. نتایج نشان می‌دهد که مدل‌های یادگیری عمیق در مواجهه با داده‌های پیچیده، عملکرد دقیق‌تر و پایدارتری نسبت به روش‌های سنتی دارند. این امر منجر به کاهش توقف‌های غیرمنتظره، بهبود سلامت تجهیزات و افزایش بهره‌وری عملیاتی در صنعت فولاد شده است.

۳-۴-۱ مطالعه موردی: پیاده‌سازی هوش مصنوعی در Tata Steel

در این بخش یک مطالعه موردی واقعی از پیاده‌سازی هوش مصنوعی در صنعت فولاد ارائه می‌شود. یکی از نمونه‌های شاخص پیاده‌سازی عملی هوش مصنوعی در صنعت فولاد، پروژهٔ پیش‌بینی نگهداری تجهیزات در کارخانه Tata Steel Shotton در بریتانیا است. در این پروژه، داده‌های واقعی حسگرهای نصب‌شده روی یاتاقان‌ها و تجهیزات کلیدی خطوط تولید جمع‌آوری و با استفاده از مدل‌های داده‌محور شامل Partial Least Squares Regression (PLSR)، شبکه‌های عصبی مصنوعی و جنگل تصادفی تحلیل شدند. نتایج نشان داد که مدل‌های مبتنی بر یادگیری

ماشین، به‌ویژه جنگل تصادفی، در پیش‌بینی سایش یاتاقان‌ها عملکرد دقیق‌تر و پایدارتری نسبت به روش‌های سنتی دارند و توانستند وقوع خرابی را پیش از رخداد تشخیص دهند. این دستاورد منجر به کاهش توقف‌های غیرمنتظره، افزایش بهره‌وری خط تولید و بهبود تصمیم‌گیری‌های نگهداری شد [۳۷].

علاوه بر این مطالعه موردی، شرکت Tata Steel در هند کارخانه (Kalinganagar) نیز با اجرای پروژه‌های هوش مصنوعی در حوزه‌های مختلف تولید از جمله پیش‌بینی دمای مذاب و بهینه‌سازی مصرف انرژی، موفق به کاهش هزینه‌های عملیاتی و بهبود کیفیت محصول شده است [۳۸].

گزارش‌های سالانه Tata Steel نیز نشان می‌دهند که این شرکت طی سال‌های اخیر بیش از ۵۵۰ مدل هوش مصنوعی را در زمینه‌هایی مانند بهره‌وری، ایمنی و کیفیت به‌کار گرفته است [۳۹]. این نمونه‌ها به‌روشنی بیانگر آن هستند که هوش مصنوعی نه تنها در سطح پژوهش، بلکه در محیط‌های صنعتی واقعی نیز به‌طور گسترده و مؤثر مورد استفاده قرار گرفته است.

۳-۵ زمان‌بندی و برنامه‌ریزی تولید

برنامه‌ریزی تولید در صنعت فولاد از مسائل پیچیده و چندهدفه است که نیازمند توازن میان سفارش‌های مشتری، زمان‌بندی تجهیزات و تأمین مواد اولیه است. الگوریتم‌های هوش مصنوعی از جمله الگوریتم‌های ژنتیک، بهینه‌سازی ازدحام ذرات و سیستم‌های ترکیبی، ابزارهای مؤثری برای بهینه‌سازی این فرآیند هستند. در مطالعه‌ای [۳۳]، یک معماری هوشمند مبتنی بر سیستم چندعاملی برای بهبود تصمیم‌گیری و زمان‌بندی تولید در کارخانه فولاد ACINOXکوبا ارائه شده است. این سیستم با ترکیب ساختارهای سلسله‌مراتبی و هم‌سطح، توانسته زمان تصمیم‌گیری را تا ۹۹/۷٪ کاهش داده و مصرف انرژی و هزینه‌ها را به شکل چشم‌گیری بهینه‌سازی کند. نتایج این رویکرد نشان می‌دهد که سیستم‌های مبتنی بر هوش مصنوعی نه‌تنها در پاسخ به شرایط متغیر و عدم قطعیت‌های موجود در تولید مؤثر هستند، بلکه به افزایش بهره‌وری و انعطاف‌پذیری در مدیریت زنجیره تولید نیز کمک می‌کنند.

۳-۶ طراحی آلیاژ و پیش‌بینی خواص محصول

با توجه به پیچیدگی خواص مکانیکی و ترکیب شیمیایی فولاد، مدل‌های یادگیری قادرند رابطه بین ترکیب عناصر و خواص نهایی محصول را استخراج و پیش‌بینی کنند [۴۰]. به‌عنوان نمونه، در مطالعه‌ای [۱۰] با استفاده از مدل یادگیری عمیق، خواص مکانیکی ورق‌های فولادی نورد گرم شامل استحکام تسلیم، استحکام کششی، درصد ازدیاد طول و انرژی ضربه بر اساس ترکیب شیمیایی و پارامترهای فرآیند تولید پیش‌بینی شد. این مدل با داده‌های واقعی صنعتی، پس از تنظیم دقیق و مقایسه با الگوریتم‌های سنتی، عملکرد بهتری از خود نشان داد و به‌صورت عملی در خطوط تولید فولاد برای پایش و کنترل آنلاین خواص مکانیکی، کاهش هزینه‌ها و بهبود کیفیت به‌کار گرفته شده است.

۳-۷ اتومات‌سازی تصمیم‌گیری در زنجیره تأمین

هوش مصنوعی با تحلیل داده‌های تقاضا، موجودی و حمل‌ونقل، فرآیند تصمیم‌گیری در زنجیره تأمین را بهینه می‌کند. شرکت‌های فولادی در چین با بهره‌گیری از الگوریتم‌های هوشمند، سفارش مواد اولیه و زمان‌بندی ارسال را مدیریت کرده‌اند که این امر موجب کاهش تأخیر، افزایش چابکی و پاسخ‌دهی سریع‌تر به تغییرات بازار شده است [۴۱].

۳-۸ کنترل زیست‌محیطی و کاهش آلاینده‌ها

هوش مصنوعی با پایش بلادرنگ انتشار CO_2 و سایر آلاینده‌ها، امکان تنظیم خودکار فرآیندها و کاهش اثرات زیست‌محیطی را فراهم می‌کند. شرکت‌های اروپایی، به ویژه در آلمان و سوئد، از سامانه‌های هوشمند برای تطبیق فرآیند تولید با استانداردهای اتحادیه اروپا استفاده می‌کنند [۵]. مطالعه‌ای [۴۲] نقش یادگیری ماشین در کاهش ردپای زیست‌محیطی و بهبود بهره‌وری تولید فولاد را بررسی کرده و نشان می‌دهد که چالش‌های داده‌ای نظیر تنوع، حجم بالا و ناپیوستگی داده‌ها با استفاده از کلان‌داده، مدل‌سازی پیشرفته و الگوریتم‌های یادگیری ماشین قابل مدیریت‌اند. این سامانه‌ها به عنوان تصمیم‌یار برای مدیران کارخانه‌ها به کار گرفته شده‌اند و مدل‌های مبتنی بر شبکه‌های عصبی مصنوعی به ویژه در شرایط دسترسی به داده‌های واقعی،

عملکرد قابل توجهی در پیش‌بینی مصرف انرژی، پایش کیفیت و کاهش ضایعات ارائه داده‌اند.

۳-۹ خلاصه کاربردها

در این بخش، خلاصه‌ای از کاربردهای عملیاتی هوش مصنوعی در صنعت فولاد ارائه می‌شود. هوش مصنوعی با بهره‌گیری از الگوریتم‌های پیشرفته یادگیری ماشین و یادگیری عمیق، در بهینه‌سازی و ارتقای فرآیندهای مختلف صنعت فولاد نقش محوری ایفا می‌کند. از بهینه‌سازی مصرف انرژی و کنترل کیفیت محصولات گرفته تا پیش‌بینی عملکرد تجهیزات، زمان‌بندی تولید، طراحی آلیاژ، اتوماتیک‌سازی زنجیره تأمین و کنترل زیست‌محیطی، کاربردهای متنوع و موثری در این صنعت دارد. این فناوری‌ها باعث افزایش بهره‌وری، کاهش هزینه‌ها، بهبود کیفیت و کاهش اثرات زیست‌محیطی می‌شوند. همچنین، استفاده از سامانه‌های هوشمند تصمیم‌یار و تحلیل‌های داده‌محور، فرآیندهای پیچیده و چندعامله تولید فولاد را به شکلی دقیق‌تر و پویا مدیریت می‌کند. نتایج این جمع‌بندی به صورت خلاصه و دسته‌بندی‌شده در جدول ۲ ارائه شده است که شامل کاربردهای عملیاتی، الگوریتم‌های رایج، نوع داده‌های مورد استفاده، اهداف اصلی و مزایای کلیدی می‌باشد.

جدول (۲): خلاصه‌ای از کاربردهای عملیاتی هوش مصنوعی در صنعت فولاد به همراه الگوریتم‌های رایج، نوع داده‌های استفاده شده، اهداف اصلی و مزایای کلیدی هر کاربرد

کاربرد عملیاتی	الگوریتم‌های رایج	نوع داده	هدف اصلی	مزایای کلیدی	مراجع
بهینه‌سازی مصرف انرژی	KNN, RF, SVR, DNN, CatBoost, LSTM	دما، فشار، ترکیب مواد	کاهش مصرف انرژی، تنظیم احتراق	صرفه‌جویی مالی، کاهش آلاینده‌ها	[۸، ۱۱، ۳۳، ۳۴]
کنترل کیفیت و تشخیص عیوب	YOLOv7, ResNet, CNN, GAN, SVM, KNN, RF	تصویر، داده نورد	شناسایی ترک، حفره، موج‌دار بودن	کاهش ضایعات، افزایش پذیرش محصول	[۱، ۲، ۳۶، ۴۳]
بهینه‌سازی فرآیند تولید	RL, SVR, RF, KNN, GAN, GNN, LSTM, DNN, Fuzzy Logic	دما، سرعت، فشار	تنظیم خودکار پارامترهای تولید	افزایش بهره‌وری، کاهش نوسانات	[۷، ۸، ۱۰، ۱۱، ۳۳، ۴۲]
نگهداری پیش‌بینانه تجهیزات	RF, SVM, LSTM, Autoencoder, GRU	ارتعاش، صوت، دما، فشار	پیش‌بینی خرابی، پایش وضعیت	کاهش توقف اضطراری، افزایش عمر تجهیز	[۴، ۵، ۴۰، ۴۴]
زمان‌بندی و برنامه‌ریزی تولید	GA, PSO, Expert System	سفارش، ظرفیت، زمان راه‌اندازی	بهینه‌سازی ترتیب تولید و زمان‌بندی	افزایش بهره‌وری، کاهش خاموشی کوره	[۳۴، ۴۳، ۴۴، ۴۵]
طراحی آلیاژ و پیش‌بینی خواص	SVR, RF, DNN, Symbolic Regression	ترکیب شیمیایی، دمای فرآیند	پیش‌بینی خواص مکانیکی	طراحی دقیق آلیاژ، کاهش آزمون‌های فیزیکی	[۱۰، ۱۵، ۴۰]
اتومات‌سازی زنجیره تأمین	RL, GA, Fuzzy Logic, CNN, Expert System	موجودی، حمل‌ونقل، تقاضا	تنظیم سفارش و ارسال	افزایش چابکی، کاهش تأخیر	[۳۶، ۴۴، ۴۶، ۴۷]
کنترل زیست‌محیطی	CNN, SVM, Fuzzy Logic, LSTM	داده گازها، دما، آلاینده‌ها و داده‌های ترکیبی	کاهش انتشار CO_2 و NOx	تطابق با استانداردها، حفظ پایداری	[۳۶، ۴۲، ۴۶، ۴۷]

۴- مزایا و دستاوردهای به‌کارگیری هوش مصنوعی در صنعت فولاد

پیاده‌سازی فناوری‌های هوش مصنوعی در صنعت فولاد تحولات چشمگیری در بهبود بهره‌وری، کاهش هزینه‌ها و ارتقای کیفیت محصولات به همراه داشته است. از جمله مهم‌ترین دستاوردها می‌توان به موارد زیر اشاره کرد:

- **افزایش بهره‌وری:** تحلیل داده‌های دقیق و پیش‌بینی‌های به موقع، امکان بهینه‌سازی پارامترهای عملیاتی را فراهم کرده و نرخ تولید را بهبود بخشیده است.

- **کاهش هزینه‌های عملیاتی:** بهینه‌سازی مصرف انرژی و کاهش ضایعات، هزینه‌های مستقیم و غیرمستقیم تولید را به طور قابل توجهی کاهش داده است.
- **ارتقای کیفیت محصولات:** استفاده از بینایی ماشین و الگوریتم‌های یادگیری عمیق، تشخیص سریع و دقیق عیوب را ممکن ساخته و به کاهش ضایعات و افزایش رضایت مشتری منجر شده است.
- **نگهداری پیش‌بینانه:** تحلیل داده‌های حسگرها، پیش‌بینی خرابی تجهیزات را تسهیل کرده و باعث کاهش توقف‌های ناگهانی و افزایش عمر مفید تجهیزات شده است.
- **افزایش انعطاف‌پذیری تولید:** مدل‌های هوش مصنوعی امکان واکنش سریع به تغییرات بازار و تقاضا را فراهم کرده و برنامه‌ریزی تولید را بهبود داده‌اند.
- **تصمیم‌گیری هوشمند:** سیستم‌های تصمیم‌یار مبتنی بر هوش مصنوعی، داده‌های بزرگ را تحلیل و به مدیران کمک می‌کنند تا تصمیمات راهبردی و عملیاتی دقیق‌تری اتخاذ کنند.
- **پایداری زیست‌محیطی:** بهینه‌سازی مصرف منابع و کاهش انتشار آلاینده‌ها، به تحقق اهداف محیط‌زیستی و تولید فولاد سبز کمک می‌کند.

این دستاوردها نشان‌دهنده نقش کلیدی هوش مصنوعی در پیشرفت پایدار و بهره‌ور صنعت فولاد است و چشم‌انداز روشنی برای گسترش کاربردهای این فناوری در آینده فراهم می‌آورد.

۵- چالش‌ها و موانع پیاده‌سازی هوش مصنوعی در صنعت فولاد

با وجود مزایای قابل‌توجه هوش مصنوعی در بهبود عملکرد صنعت فولاد، استقرار موفق آن با چالش‌هایی در سطوح فنی، سازمانی و اجرایی همراه است. این بخش به مهم‌ترین موانع پیش‌رو در محیط‌های تولید فولاد اشاره دارد:

۵-۱ کیفیت و دسترسی به داده‌ها

یکی از موانع اساسی، نبود داده‌های دقیق، کامل و ساخت‌یافته در مراحل مختلف فرآیندهای فولادسازی است. پیچیدگی تولید، ضعف زیرساخت‌های اندازه‌گیری و فقدان استانداردهای یکپارچه برای ذخیره‌سازی داده، موجب کاهش دقت مدل‌های یادگیری ماشین می‌شود.

دلیل اهمیت در فولاد: داده‌های نامطمئن یا ناقص می‌توانند پیش‌بینی فرآیندها و کنترل کیفیت محصولات فولادی را دچار خطا کنند.

۵-۲ پیچیدگی مدل‌سازی فرایندها

ماهیت چندمتغیره، پویا و غیرخطی فرایندهای تولید فولاد، مدل‌سازی دقیق را دشوار می‌سازد. اجرای مؤثر الگوریتم‌های هوش مصنوعی مستلزم دانش میان‌رشته‌ای در متالورژی، مهندسی فرایند و یادگیری ماشین است. کمبود نیروی متخصص یکی از محدودیت‌های اصلی است.

دلیل اهمیت در فولاد: پیچیدگی فرآیندهای فولادسازی باعث می‌شود مدل‌های هوش مصنوعی مستقیماً بر کیفیت محصول و بهره‌وری خطوط تولید تأثیر بگذارند.

۵-۳ ناسازگاری با زیرساخت‌های موجود

بسیاری از کارخانه‌های فولاد از سیستم‌های کنترل قدیمی و غیرهوشمند استفاده می‌کنند که با فناوری‌های نوین سازگار نیستند. فقدان حسگرهای هوشمند، اتصال ضعیف سامانه‌ها و مشکلات امنیت سایبری، استقرار راهکارهای هوش مصنوعی را با چالش مواجه می‌کند.

دلیل اهمیت در فولاد: عدم تطابق زیرساخت‌ها با سیستم‌های هوشمند می‌تواند موجب اختلال در خطوط تولید و افزایش هزینه‌های عملیاتی شود.

۵-۴ مقاومت فرهنگی و سازمانی

پذیرش فناوری‌های نوین در صنعت فولاد نیازمند تغییر فرهنگ سازمانی و توانمندسازی کارکنان است. نگرانی از جایگزینی شغلی و ناآشنایی با مزایای فناوری، مقاومت در برابر تغییر را افزایش می‌دهد.

دلیل اهمیت در فولاد: مشارکت فعال کارکنان و آموزش هدفمند، موفقیت پروژه‌های هوشمندسازی و حفظ بهره‌وری در محیط‌های تولید فولاد را تضمین می‌کند.

۵-۵ تفسیرپذیری و اعتماد به مدل‌ها

مدل‌های یادگیری عمیق پیچیده و غیرقابل تفسیر هستند. در صنعت فولاد که تصمیم‌ها باید مستند و دقیق باشند، این موضوع موجب تردید در اعتماد به مدل‌ها می‌شود.

دلیل اهمیت در فولاد: عدم شفافیت مدل‌ها می‌تواند مانع تصمیم‌گیری سریع و بهینه در فرآیندهای تولید فولاد شود.

۵-۶ امنیت سایبری و حریم خصوصی داده‌ها

استفاده از اینترنت اشیا و جمع‌آوری داده‌های حساس تولید، خطرات امنیتی جدی دارد. حملات سایبری به سیستم‌های کنترل تولید فولاد می‌توانند پیامدهای گسترده‌ای داشته باشند.

دلیل اهمیت در فولاد: حفاظت از داده‌ها و زیرساخت‌ها برای جلوگیری از توقف تولید یا آسیب به تجهیزات حیاتی الزامی است.

۵-۷ هزینه و بازگشت سرمایه

هوشمندسازی نیازمند سرمایه‌گذاری قابل‌توجه در سخت‌افزار، نرم‌افزار و آموزش است. محاسبه دقیق بازگشت سرمایه و استفاده از رویکردهای مرحله‌ای و چابک، برای کاهش ریسک‌های اقتصادی و فنی حیاتی است.

دلیل اهمیت در فولاد: سرمایه‌گذاری نادرست می‌تواند موجب کاهش رقابت‌پذیری و بهره‌وری کلان صنعت فولاد شود.

۶- آینده‌پژوهی و روندهای نوظهور در کاربرد هوش مصنوعی در صنعت فولاد

تحول دیجیتال و پیشرفت سریع فناوری‌های نوین، چشم‌اندازهای تازه‌ای را برای صنعت فولاد ترسیم کرده است. هوش مصنوعی، به‌عنوان پیشران اصلی این تغییرات، در سال‌های آینده نقشی حیاتی در افزایش خودکارسازی، بهره‌وری و پایداری ایفا خواهد کرد. روندهای کلیدی آینده در به‌کارگیری هوش مصنوعی در صنعت فولاد بر اساس سطح بلوغ فناوری و میزان تأثیر بالقوه، در شکل ۲ نمایش داده شده‌اند. در ادامه، شرح مختصری از هر یک از این روندها ارائه می‌شود:

شکل (۲): روندهای کلیدی آینده در به‌کارگیری هوش مصنوعی در صنعت فولاد

۶-۱ حرکت به‌سوی کارخانه‌های هوشمند و خودران

یکی از افق‌های بلندمدت صنعت فولاد، ایجاد کارخانه‌هایی است که با حداقل مداخله انسانی، از طریق ترکیب هوش مصنوعی، رباتیک، اینترنت اشیاء صنعتی و فناوری‌های لبه‌محور بتوانند به‌صورت کاملاً خودکار، پویا و واکنش‌گرا عمل کنند. در این چشم‌انداز، الگوریتم‌های پیش‌بینی، کنترل‌کننده‌های هوشمند و سیستم‌های تصمیم‌یار در تمامی مراحل، از شارژ مواد خام تا بسته‌بندی محصول نهایی، نقش اصلی را ایفا خواهند کرد.

۶-۲ ترکیب هوش مصنوعی با فناوری‌های نوظهور

یکپارچه‌سازی هوش مصنوعی با فناوری‌هایی مانند دوقلوهای دیجیتال، یادگیری فدرال، بلاک‌چین و واقعیت افزوده، امکان تحلیل، شبیه‌سازی و بهینه‌سازی پیشرفته‌تری از فرآیندها را فراهم خواهد کرد.

۶-۳ توسعه مدل‌های سبک و قابل‌تفسیر

افزایش تقاضا برای تصمیم‌گیری‌های شفاف و مستند در صنعت فولاد، جهت‌گیری پژوهش‌ها را به‌سمت مدل‌های قابل‌تفسیر و کم‌هزینه با دقت بالا سوق داده است.

۶-۴ گسترش تصمیم‌یارهای هوشمند در زنجیره تأمین

در آینده، هوش مصنوعی فراتر از خطوط تولید، در بهینه‌سازی لجستیک، پیش‌بینی تقاضا و برنامه‌ریزی حمل‌ونقل نقش فزاینده‌ای

ایفا خواهد کرد، به‌ویژه از طریق مدل‌های یادگیری تقویتی و تحلیل عدم قطعیت.

۶-۵ تولید پایدار و زیست‌سازگار با کمک هوش مصنوعی

مدل‌های هوشمند قادر خواهند بود بهینه‌سازی انرژی، کاهش انتشار آلاینده‌ها و مدیریت مؤثر پسماند را در راستای اهداف زیست‌محیطی جهانی تحقق بخشند.

۶-۶ نقش رو به رشد هوش مصنوعی مولد

با پیشرفت مدل‌های هوش مصنوعی مولد مانند GPT در حوزه‌های صنعتی، در آینده، هوش مصنوعی قادر خواهد بود به‌صورت خودکار طرح‌های مهندسی پیشنهاد دهد، سناریوهای بهینه طراحی کند و حتی راهکارهای نوآورانه برای مسائل فرآیندی خلق کند. این روند، خلاقیت انسانی را تقویت کرده و بهره‌وری طراحی را ارتقاء خواهد داد.

۶-۷ خودکارسازی تصمیم‌گیری در خطوط تولید

دستیابی به کنترل کامل و هوشمند فرآیندهای تولید، از طریق توسعه الگوریتم‌های پیشرفته کنترل، تشخیص سریع شرایط بحرانی و تضمین ایمنی، از اهداف نهایی آینده‌پژوهی در این حوزه است.

۷- نتیجه‌گیری

صنعت فولاد به‌عنوان یکی از بخش‌های کلیدی اقتصاد جهانی، در مسیر گذار به تولید هوشمند، به‌طور فزاینده‌ای به فناوری‌های هوش مصنوعی متکی شده است. مطالعات اخیر نشان می‌دهد که هوش مصنوعی با بهینه‌سازی فرآیندها، پیش‌بینی خرابی‌ها، کنترل کیفیت و مدیریت منابع، نقش مؤثری در افزایش بهره‌وری و پایداری عملیاتی ایفا کرده است. الگوریتم‌های یادگیری ماشین، یادگیری عمیق و بینایی ماشین در کاربردهایی مانند تشخیص عیوب، کاهش مصرف انرژی و طراحی آلیاژهای جدید مورد استفاده قرار گرفته‌اند و موجب بهبود تصمیم‌گیری و انعطاف‌پذیری تولید شده‌اند. با این حال، چالش‌هایی مانند نیاز به داده‌های دقیق، نبود زیرساخت‌های مناسب، مسائل امنیتی و تفسیرناپذیری برخی مدل‌ها، مانع پیاده‌سازی گسترده این فناوری‌ها هستند. همچنین، رعایت الزامات اخلاقی، شفافیت الگوریتم‌ها و ملاحظات زیست‌محیطی باید در آینده مورد توجه ویژه قرار گیرد.

برای مدیران صنعت فولاد، مسیر عملیاتی تحقق چشم‌انداز هوشمندسازی به شرح زیر پیشنهاد می‌شود:

- سرمایه‌گذاری راهبردی در زیرساخت‌های دیجیتال و تربیت نیروی انسانی متخصص: ایجاد بسترهای داده‌محور، تجهیز خطوط تولید به حسگرها و سیستم‌های هوشمند، و آموزش نیروهای متخصص برای بهره‌برداری مؤثر از هوش مصنوعی
- تقویت همکاری بین صنایع، مراکز پژوهشی و شرکت‌های فناور: اجرای پروژه‌های مشترک تحقیق و توسعه، ایجاد شبکه‌های دانش و اشتراک‌گذاری داده‌های صنعتی برای افزایش سرعت نوآوری
- پیاده‌سازی راهکارهای هوشمند با رویکرد پایداری: کاهش مصرف انرژی و آلاینده‌ها، مدیریت مؤثر پسماند و ارتقای شاخص‌های زیست‌محیطی در فرآیندهای تولید
- استفاده تدریجی و هدفمند از فناوری‌های نوظهور: برنامه‌ریزی برای ادغام رباتیک، اینترنت اشیاء صنعتی، مدل‌های پیش‌بینی و تصمیم‌یار هوشمند در فرآیندهای تولید و زنجیره تأمین

در نهایت، آینده صنعت فولاد به توانایی مدیران در ادغام هوشمندانه فناوری‌های نوین با فرآیندهای تولید، مدیریت و زیست‌محیطی وابسته است و مسیر مشخصی برای تحقق کارخانه‌های خودران، زنجیره تأمین هوشمند و تولید پایدار ارائه می‌دهد.

References

[1] R. Ma, J. Chen, Y. Feng, Z. Zhou, and J. Xie, "ELA-YOLO: An efficient method with linear attention for steel surface defect detection during manufacturing," Adv. Eng. Inform., vol. 65, p. 103377, 2025.

[2] V. Nath, C. Chattopadhyay, and K. A. Desai, "NSLNet: An improved deep learning model for steel surface defect classification utilizing small training datasets," Manuf. Lett., vol. 35, pp. 39–42, 2023.

[3] W. Fang, J. X. Huang, T. X. Peng, Y. Long, and F. X. Yin, "Machine learning-based performance predictions for steels considering manufacturing process parameters: A review," J. Iron Steel Res. Int., vol. 31, no. 7, pp. 1555–1581, 2024.

[4] J. Jakubowski, N. Wojak-Strzelecka, R. P. Ribeiro, S. Pashami, S. Bobek, J. Gama, and G. J. Nalepa, "Artificial intelligence approaches for predictive maintenance in the steel industry: A survey," arXiv preprint arXiv:2405.12785, 2024.

[5] C. Pepe, G. Farella, G. Bartucci, and S. M. Zanoli, "Recent innovations in computer and automation engineering for performance improvement in the steel industry production chain: A review," Energies, vol. 18, no. 8, 2025.

[6] D. Zhou, K. Xu, Z. Lv, J. Yang, M. Li, F. He, and G. Xu, "Intelligent manufacturing technology in the steel industry of China: A review," Sensors, vol. 22, no. 21, p. 8194, 2022.

[7] K. Tsutsui, T. Namba, K. Kihara, J. Hirata, S. Matsuo, and K. Ito, "Current trends on deep learning techniques applied in iron and steel making field: A review," ISIJ Int., vol. 64, no. 11, pp. 1619–1640, 2024.

[8] Redchuk and F. Walas Mateo, "New business models on artificial intelligence—the case of the optimization of a blast furnace in the steel industry by a machine learning solution," Appl. Syst. Innov., vol. 5, no. 1, p. 6, 2021.

[9] D. Cemernek, S. Cemernek, H. Gursch, A. Pandeshwar, T. Leitner, M. Berger, G. Klösch, and R. Kern, "Machine learning in continuous casting of steel: A state-of-the-art survey," J. Intell. Manuf., vol. 33, no. 6, pp. 1561–1579, 2022.

[10] Q. Xie, M. Suvarna, J. Li, X. Zhu, J. Cai, and X. Wang, "Online prediction of mechanical properties of hot rolled steel plate using machine learning," Mater. Des., vol. 197, p. 109201, 2021.

[11] S. W. Choi, B. G. Seo, and E. B. Lee, "Machine learning-based tap temperature prediction and control for optimized power consumption in stainless electric arc furnaces (EAF) of steel plants," Sustainability, vol. 15, no. 8, p. 6393, 2023.

[12] M. Khaledi, A. Rashidi Mehrabadi, and M. Mirabi, "Prediction of the corrosion and scaling index of industrial cooling water circulation circuits using artificial intelligence: Case study of circulating water circuits in electric arc furnaces of Khuzestan Steel," SSRN, 2024.

[13] Y. Zhang, C. J. Zhang, K. Zeng, L. Zhu, and Y. Han, "Research on terminal control model of intelligent mining of flame spectral information of converter mouth in late smelting stage," Ironmaking Steelmaking, vol. 48, no. 6, pp. 677–684, 2021.

[14] S. Li, S. Wang, W. Li, G. Zhang, Z. Huang, X. Luo, and H. Yu, "Rolled thickness prediction for titanium/steel-clad plates based on combined method of theoretical and neural network," Steel Res. Int., vol. 96, no. 3, p. 2400602, 2025.

[15] S. Y. Lee, B. A. Tama, S. J. Moon, and S. Lee, "Steel surface defect diagnostics using deep convolutional neural network and class activation map," Appl. Sci., vol. 9, no. 24, p. 5449, 2019.

[16] K. Choi, K. Koo, and J. S. Lee, "Development of defect classification algorithm for POSCO rolling strip surface inspection system," in Proc. 2006 SICE–ICASE Int. Joint Conf., Oct. 2006, pp. 2499–2502. IEEE.

[17] L. North, K. Blackmore, K. Nesbitt, and M. R. Mahoney, "Methods of coke quality prediction: A review," Fuel, vol. 219, pp. 426–445, 2018.

[18] C. Yang, C. Yang, J. Li, Y. Li, and F. Yan, "Forecasting of iron ore sintering quality index: A latent variable method with deep inner structure," Comput. Ind., vol. 141, p. 103713, 2022.

[19] S. Liu, X. Liu, Q. Lyu, and F. Li, "Comprehensive system based on a DNN and LSTM for predicting sinter composition," Appl. Soft Comput., vol. 95, p. 106574, 2020.

[20] F. He and L. Zhang, "Mold breakout prediction in slab continuous casting based on combined method of GA–BP neural network and logic rules," Int. J. Adv. Manuf. Technol., vol. 95, no. 9, pp. 4081–4089, 2018.

[21] J. Cheng, C. Zhao-Zhen, T. Nai-Biao, Y. Ji-Lin, and Z. Miao-Yong, "Molten steel breakout prediction based on genetic algorithm and BP

neural network in continuous casting process," in Proc. 31st Chin. Control Conf., Jul. 2012, pp. 3402–3406. IEEE.
[22] W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng, "A new steel defect detection algorithm based on deep learning," Comput. Intell. Neurosci., vol. 2021, no. 1, p. 5592878, 2021.
[23] B. Si, M. Yasengjiang, and H. Wu, "Deep learning-based defect detection for hot-rolled strip steel," in J. Phys.: Conf. Ser., vol. 2246, no. 1, p. 012073. IOP Publishing, 2022.
[24] J. Jang, D. Van, H. Jang, D. H. Baik, S. D. Yoo, J. Park, S. Mhin, J. Mazumder, and S. H. Lee, "Residual neural network-based fully convolutional network for microstructure segmentation," Sci. Technol. Weld. Join., vol. 25, no. 4, pp. 282–289, 2020.
[25] B. Zhu, Z. Chen, F. Hu, X. Dai, L. Wang, and Y. Zhang, "Feature extraction and microstructural classification of hot stamping ultra-high strength steel by machine learning," JOM, vol. 74, no. 9, pp. 3466–3477, 2022.
[26] "Accelerating phase-field predictions via recurrent neural networks learning the microstructure evolution in latent space," [Unpublished/Incomplete Citation].
[27] H. Di, X. Ke, Z. Peng, and D. Dongdong, "Surface defect classification of steels with a new semi-supervised learning method," Opt. Lasers Eng., vol. 117, pp. 40–48, 2019.
[28] Y. Gao, H. Zhang, L. Zhu, F. Xie, and D. Xiao, "ZFD-Net: Zinc flower defect detection model of galvanized steel surface based on improved YOLOv5," PLoS One, vol. 20, no. 6, p. e0325507, 2025.
[29] B. Kim, J. Kwon, S. Choi, J. Noh, K. Lee, and J. Yang, "Corrosion image monitoring of steel plate by using k-means clustering," J. Korean Inst. Surf. Eng., vol. 54, no. 5, pp. 278–284, 2021.
[30] R. Zheng, Y. Bao, L. Zhao, and L. Xing, "Prediction of steelmaking process variables using K-medoids and a time-aware LSTM network," Heliyon, vol. 10, no. 12, 2024.
[31] Y. Ji, S. Liu, M. Zhou, Z. Zhao, X. Guo, and L. Qi, "A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems," Inf. Sci., vol. 589, pp. 360–375, 2022.
[32] C. Song, J. Cao, L. Wang, J. Xiao, and Q. Zhao, "Transverse thickness profile control of electrical steel in 6-high cold rolling mills based on the GA–PSO hybrid algorithm," Int. J. Adv. Manuf. Technol., vol. 121, no. 1, pp. 295–308, 2022.
[33] R. Ricardo Rodríguez, I. F. Benítez, G. González Yero, and J. R. Núñez Alvarez, "Multi-agent system for steel manufacturing process," Int. J. Electr. Comput. Eng. (IJECE), vol. 12, no. 3, pp. 2441–2453, 2022.
[34] K. Kerdprasop, N. Kerdprasop, and P. Chuaybamroong, "Deep learning and machine learning models to predict energy consumption in steel industry," Int. J. Mach. Learn., vol. 13, pp. 142–145, 2023.
[35] X. Chen, J. Van Hillegersberg, E. Topan, S. Smith, and M. Roberts, "Application of data-driven models to predictive maintenance: Bearing wear prediction at TATA steel," Expert Syst. Appl., vol. 186, p. 115699, 2021.
[36] U. Samal, "Evolution of machine learning and deep learning in intelligent manufacturing: A bibliometric study," Int. J. Syst. Assur. Eng. Manag., pp. 1–17, 2025.
[37] X. Chen, J. Van Hillegersberg, E. Topan, S. Smith, and M. Roberts, "Application of data-driven models to predictive maintenance: Bearing wear prediction at TATA steel," Expert Syst. Appl., vol. 186, p. 115699, 2021.
[38] McKinsey & Company, "How a steel plant in India tapped the value of data and won global acclaim," 2020.
[39] Tata Steel, Integrated Report 2023–24: Intellectual Capital – Industry 4.0, Tata Steel Official Reports, 2024.
[40] R. Kumar Balaraman, S. Hussain, J. K. Ong, Q. Y. Tan, and N. Raghavan, "Feature-driven density prediction of maraging steel additively manufactured samples using pyrometer sensor and supervised machine learning," IEEE Access, vol. 12, pp. 172892–172909, 2024.
[41] Khalili-Fard, F. Sabouhi, and A. Bozorgi-Amiri, "Data-driven robust optimization for a sustainable steel supply chain network design: Toward the circular economy," Comput. Ind. Eng., vol. 195, p. 110408, 2024.
[42] V. Colla, C. Pietrosanti, E. Malfa, and K. Peters, "Environment 4.0: How digitalization and machine learning can improve the environmental footprint of the steel production processes," Matér. Tech., vol. 108, no. 5–6, p. 507, 2020.
[43] H. Zhao, Y. Chen, B. Dang, and X. Jian, "Research on steel production scheduling optimization based on deep learning," in Proc. 4th Int. Symp. Artif. Intell. Intell. Manuf. (AIIM), Dec. 2024, pp. 813–816. IEEE.

[44] M. Coniglio, A. Cimino, and V. Corvello, “Artificial intelligence and the future of supply chain management,” in Int. Res. Innov. Forum, Cham: Springer Nature Switzerland, 2024, pp. 43–52.

[45] Ş. Duymaz and A. F. Güneri, “The application of machine learning algorithms in the estimation of production lead times: A case study of a steel construction manufacturing company,” J. Adv. Manuf. Eng. (JAME), vol. 5, no. 1, 2024.

[46] Khalili-Fard, F. Sabouhi, and A. Bozorgi-Amiri, “Data-driven robust optimization for a sustainable steel supply chain network design: Toward the circular economy,” Comput. Ind. Eng., vol. 195, p. 110408, 2024.

[47] T.-L. Nguyen, P.-H. Nguyen, H.-A. Pham, T.-G. Nguyen, D.-T. Nguyen, T.-H. Tran, H.-C. Le, and H.-T. Phung, “A novel integrating data envelopment analysis and spherical fuzzy MCDM approach for sustainable supplier selection in steel industry,” Mathematics, vol. 10, p. 1897, 2022.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 1 - 16

Smart Steel Industry: A Review of Practical Applications of Artificial Intelligence in Optimization, Sustainability, and Digital Transformation of Production Processes

Razieh Sheikhpour*

Associate Professor, Department of Computer Engineering, Faculty of Engineering, Ardakan University, Ardakan, Iran

Article Information

Original Research Paper

Received:
2025 July 06

Accepted:
2025 October 12

Keywords:
Artificial Intelligence, Steel Industry, Machine Learning, Process Optimization, Digital Transformation

Corresponding Author*:
rsheikhpour@ardakan.ac.ir

Abstract

The Steel industry, as one of the vital pillars of the global economy, faces challenges such as market fluctuations, rising energy costs, the necessity to reduce pollutants, and increasing competition in productivity. In this context, artificial intelligence (AI) serves as a key driver of digital transformation, playing a significant role in the optimization and intelligent automation of production processes. This review explores the operational applications of AI in the steel industry. It begins by introducing the fundamental concepts of industrial intelligence and AI algorithms, followed by an examination of various AI applications such as energy consumption optimization, quality control, predictive maintenance, material property prediction, supply chain decision-making, and production line automation, along with the benefits of AI implementation. Furthermore, the challenges of AI deployment, including the need for accurate data, technological infrastructure, security concerns, and model interpretability, are analyzed. Finally, emerging trends, such as the integration of AI with technologies like digital twins, federated learning, and interpretable models, are introduced. The findings of this review indicate that the targeted use of AI not only enhances productivity and reduces costs but also paves the way for achieving sustainable and responsive production.

: 10.22034/ABMIR.2025.23461.1146

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم پاییز و زمستان ۱۴۰۴، صفحات۳۳ - ۱۷

پیش‌بینی سری‌های زمانی قیمت فولاد با استفاده از معماری ترنسفورمر پچ‌محور

فاطمه زارع مهرجردی[۱*]، عباس نوروزی[۲]

[۱] استادیار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران

[۲] کارشناسی ارشد، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر، دانشگاه میبد، میبد، ایران

چکیده

مقاله پژوهشی

فولاد به عنوان یکی از مهم‌ترین مصالح در صنایع مختلف، نقشی حیاتی در اقتصاد جهانی ایفا می‌کند. اما نوسانات بالای قیمت، وابستگی‌های پیچیده و غیرخطی و عدم توانایی مدل‌های سنتی در ثبت همبستگی‌های بلندمدت، پیش‌بینی دقیق آن را به یک چالش جدی تبدیل کرده است. این پژوهش برای غلبه بر این نارسایی‌ها، یک رویکرد جامع برای پیش‌بینی سری‌های زمانی قیمت فولاد با استفاده از مدل‌های یادگیری عمیق ارائه می‌دهد که شامل سه مرحله اصلی است: در مرحله اول، پیش‌پردازش داده‌ها از قبیل حذف مقادیر گم‌شده، نرمال‌سازی و ایجاد پنجره‌های زمانی ترتیبی انجام‌شده است؛ در مرحله دوم، مدل‌های یادگیری عمیق، شامل مدل‌های شبکه‌های عصبی بازگشتی، شبکه‌های ترنسفورمر (PatchTST) و شبکه‌های کانولوشنی آموزش داده‌شده‌اند و درنهایت، مرحله سوم ارزیابی عملکرد و مقایسه مدل‌ها با استفاده از معیارهای RMSE، MAE و R^2 صورت گرفته است. نتایج ارزیابی نشان می‌دهد که مدل PatchTST با بهره‌گیری از مکانیزم توجه و پردازش داده‌ها به صورت قطعات یا پچ توانسته است وابستگی‌های پیچیده و بلندمدت را با دقت بالاتری نسبت به سایر مدل‌ها تشخیص دهد؛ به‌طور خاص، PatchTST در پیش‌بینی قیمت‌های فولاد توانست به دقت پیشرو R^2 با مقدار ۰/۹۸۱۵ و حداقل خطای RMSE با مقدار ۰/۰۳۰۴ دست یابد. این در حالی است که مدل‌های RNN به دلیل ماهیت ترتیبی و محدودیت‌های ساختاری خود، عملکرد ضعیف‌تری ارائه دادند. این یافته‌ها بر برتری مدل‌های مبتنی بر ترنسفورمر در پیش‌بینی دقیق سری‌های زمانی پیچیده تاکید می‌کند.

تاریخ دریافت:
۱۴۰۴/۰۵/۲۵
تاریخ پذیرش:
۱۴۰۴/۰۸/۱۳

کلیدواژه‌ها:
پیش‌بینی قیمت فولاد، یادگیری عمیق، سری زمانی، شبکه‌های عصبی بازگشتی، شبکه ترنسفورمر

نویسنده‌مسئول:
fzare@meybod.ac.ir

: 10.22034/ABMIR.2025.23536.1152

E-ISSN: 2821-2037

۱- مقدمه

فولاد به عنوان یکی از مهم‌ترین و پرمصرف‌ترین مواد اولیه در جهان، نقشی حیاتی در صنایع گوناگون ازجمله ساخت‌وساز، خودروسازی، زیرساخت‌ها و تولید ماشین‌آلات ایفا می‌کند. به دلیل اهمیت استراتژیک این محصول، نوسانات قیمت آن تأثیر مستقیمی بر اقتصاد جهانی و تصمیمات تجاری شرکت‌ها دارد. از این رو، پیش‌بینی دقیق و به‌موقع قیمت فولاد به منظور مدیریت ریسک، برنامه‌ریزی تولید و بهینه‌سازی زنجیره تأمین، از چالش‌های کلیدی در مدیریت صنعتی و مالی به شمار می‌رود. بر اساس مطالعات انجام شده، جهت پیش‌بینی سری‌های زمانی قیمت‌ها، دو رویکرد اصلی روش‌های سنتی و روش‌های یادگیری عمیق وجود دارد.

در طول دهه‌های گذشته، روش‌های آماری و اقتصادسنجی متعددی برای پیش‌بینی قیمت‌ها در بازارهای مالی ارائه شده‌اند که به عنوان روش‌های سنتی شناخته می‌شوند. مدل‌های ARIMA[1] به عنوان یکی از اولین و پرکاربردترین رویکردهای سنتی پیش‌بینی سری‌های زمانی است. این مدل ترکیبی از سه جز اصلی خودرگرسیون، تفاضل‌گیری و میانگین متحرک است. مدل ARIMA با استفاده از این سه جز، می‌تواند الگوهای خطی موجود در داده‌های سری زمانی را شناسایی کرده و برای پیش‌بینی مقادیر آینده استفاده کند. از آنجایی که قیمت فولاد تحت تأثیر عوامل پیچیده‌ای هم‌چون تغییرات عرضه و تقاضا، شرایط کلان اقتصادی، سیاست‌های تجاری، و حتی تحولات ژئوپلیتیکی قرار دارد باعث شده پیش‌بینی آن با استفاده از روش‌های سنتی آماری مانند ARIMA دشوار باشد [۱].

برای تحلیل دقیق‌تر نوسانات بازار، مدل‌های خانواده GARCH[2] [۲]توسعه یافتند تا با در نظر گرفتن واریانس ناهمسان شرطی، پیش‌بینی ریسک را بهبود بخشند. هم‌چنین، رویکردهای چندمتغیره مانند VAR[3] [۳]به محققان اجازه دادند تا روابط متقابل بین قیمت کالا و سایر متغیرهای کلان اقتصادی را بررسی کنند. با این حال، همان‌طور که در پژوهش‌های اخیر اشاره شده، این مدل‌های سنتی اغلب در ثبت الگوهای غیرخطی و پیچیده در سری‌های زمانی مالی با محدودیت‌هایی مواجه هستند.

در سال‌های اخیر با پیشرفت تکنیک‌های یادگیری عمیق، مدل‌های مبتنی بر شبکه‌های عصبی توانستند با یادگیری خودکار ویژگی‌ها و کشف الگوهای غیرخطی از داده‌ها در پیش‌بینی سری‌های زمانی مالی موفق ظاهر شوند [٤]. شبکه‌های عصبی بازگشتی[4] به دلیل ساختار خود، توانایی ذاتی در پردازش داده‌های ترتیبی دارند. مدل‌های پیشرفته‌تر مانند شبکه‌های حافظه بلندمدت کوتاه[5] (LSTM)و واحدهای بازگشتی دروازه‌ای[6] (GRU)، با حل مشکل محو شدن گرادیان[7] در RNN های ساده، توانستند وابستگی‌های بلندمدت در داده‌ها را به شکل بهتری مدل‌سازی کنند. به همین دلیل، LSTMو GRU به طور گسترده در پیش‌بینی قیمت کالاها و سهام مورد استفاده قرارگرفته‌اند [٥, ٦].

در سال‌های اخیر، معماری ترنسفورمر[8] با معرفی مکانیزم توجه به خود[9] انقلابی در پردازش توالی‌ها ایجاد کرد. این معماری به مدل اجازه می‌دهد تا به جای پردازش ترتیبی، به صورت موازی به تمام نقاط ورودی توجه کند و روابط پیچیده و غیرخطی بین آن‌ها را بدون از دست دادن اطلاعات، کشف کند. مدل PatchTST یک پیاده‌سازی نوین از ترنسفورمر است که به طور خاص برای سری‌های زمانی طراحی‌شده است و با تقسیم داده‌ها به قطعات[10]، توانایی ترنسفورمر را در یادگیری الگوهای محلی و بلندمدت به شکل چشم‌گیری بهبود می‌بخشد [٤, ٧].

با توجه به روش‌های مختلف پیش‌بینی سری زمانی مبتنی بر یادگیری عمیق، نیاز به یک تحلیل مقایسه‌ای جامع و کمی بین معماری‌های مختلف شبکه‌های عصبی بازگشتی، ترنسفورمرها و شبکه‌های کانولوشنی برای پیش‌بینی قیمت فولاد احساس می‌شود. تحلیل پژوهش‌های پیشین نشان می‌دهد که علی‌رغم اهمیت حیاتی

[1] Autoregressive Integrated Moving Average (ARIMA)
[2] Generalized Autoregressive Conditional Heteroskedasticity (GARCH)
[3] Vector Autoregression (VAR)
[4] Recurrent Neural Network (RNN)
[5] Long Short-Term Memory (LSTM)
[6] Gated Recurrent Unit (GRU)
[7] Vanishing Gradient
[8] Transformer
[9] Self-Attention
[10] Patches

پیش‌بینی قیمت فولاد، سه نارسایی کلیدی در ادبیات وجود دارد که توجیهی برای پژوهش حاضر محسوب می‌شود: نخست، کمبود یک مقایسه جامع و نظام‌مند بین نسل‌های مختلف مدل‌های یادگیری عمیق نوین (ازجمله مدل‌های بازگشتی، کانولوشنی و ترنسفورمر) در حوزه قیمت فولاد مشاهده می‌شود؛ تحقیقات پیشین اغلب بر روی یک یا دو مدل خاص متمرکز بوده‌اند. دوم، مدل‌های بازگشتی سنتی(RNN/LSTM) ، که رایج‌ترین مدل‌های یادگیری عمیق در این حوزه هستند، در مواجهه با وابستگی‌های طولانی، غیرخطی و نوسانات شدید قیمت فولاد، دچار افت شدید دقت و پدیده محوشدگی گرادیان می‌شوند. سوم، علیرغم موفقیت‌های چشم‌گیر معماری‌های پیشرفته ترنسفورمر مانند PatchTST در سری‌های زمانی عمومی، کاربرد و ارزیابی عمیق این مدل‌های جدید در داده‌های تخصصی و پیچیده قیمت فولاد به‌طور قابل‌توجهی کم است. بنابراین، این پژوهش با هدف پر کردن این شکاف‌ها، یک ارزیابی مقایسه‌ای دقیق بین معماری‌های RNN، TCN و PatchTST را بر روی داده‌های واقعی قیمت فولاد انجام می‌دهد تا برترین معماری برای پیش‌بینی‌های استراتژیک این حوزه را مشخص سازد. بنابراین هدف از این پژوهش، فراتر از یک مقایسه ساده، ارزیابی دقیق عملکرد مدل‌های پیشرفته، به خصوص مدل PatchTST، در مسئله چالش‌برانگیز پیش‌بینی قیمت فولاد است. ما با استفاده از یک مجموعه داده واحد، نه تنها کارآمدترین رویکرد را شناسایی می‌کنیم، بلکه در پی ارائه یک بینش ساختاری هستیم تا مشخص شود که چگونه مزیت‌های معماری ترنسفورمر، مانند مکانیزم پچ و توجه[1]، آن را قادر می‌سازد تا وابستگی‌های پیچیده و بلندمدت قیمت فولاد را با دقت بالاتری نسبت به مدل‌های بازگشتی سنتی مدل‌سازی کند و یک معیار عملکرد جدید در این حوزه تعیین نماید.

ساختار ادامه این مقاله به شرح زیر است .در بخش دوم، مروری بر کارهای پیشین انجام شده در حوزه پیش‌بینی سری زمانی قیمت‌ها با تمرکز بر مدل‌های یادگیری عمیق ارائه می‌گردد .بخش سوم به شرح جزئیات روش پیشنهادی می‌پردازد و مجموعه داده، مدل‌های انتخابی و تنظیمات آموزش را معرفی می‌کند. در بخش چهارم، نتایج حاصل از ارزیابی‌های کمی و تحلیل‌های مقایسه‌ای ارائه شده و برتری ساختاری مدلPatchTST مورد بحث قرار می‌گیرد. درنهایت، بخش پنجم به جمع‌بندی یافته‌های اصلی پژوهش و ارائه پیشنهادهایی برای تحقیقات آتی اختصاص دارد.

۲- مروری بر کارهای گذشته

پیش‌بینی دقیق قیمت کالاها، ازجمله فولاد، سال‌هاست که موضوع پژوهش‌های فراوانی بوده است. با توجه به مطالعات انجام شده در این بخش مروری بر تعدادی مقالات انجام شده در پیش‌بینی سری‌های زمانی با روش‌های سنتی و یادگیری عمیق انجام شده است. لازم به ذکر است، از آنجایی که بازارهای کالایی مانند فولاد، نفت، طلا و رمزارزها همگی دارای ویژگی‌های دینامیکی مشترکی مانند نوسان‌پذیری[2] بالا، وابستگی‌های غیرخطی پیچیده و ماهیت پرهرج و مرج هستند، متدولوژی‌های موفق در پیش‌بینی قیمت یک کالا معمولاً قابل‌تعمیم به کالاهای دیگر نیز می‌باشند. بنابراین، بررسی مطالعات انجام شده در حوزه‌های مشابه مانند طلا، نفت خام و بازارهای سهام، برای شناسایی کارآمدترین مدل‌سازی برای چنین دینامیک‌های پیچیده‌ای، ضروری و قابل قبول است.

۲-۱ پیش‌بینی سری‌های زمانی با روش‌های سنتی

گوها و همکارانش [۸] از رویکرد سنتی ARIMA برای پیش‌بینی قیمت طلا در کشور هند استفاده کرده‌اند. مطالعه آن‌ها بر اساس داده‌های ماهانه از نوامبر ۲۰۰۳ تا ژانویه ۲۰۱۴ بوده و پس از مقایسه چندین ترکیب مختلف پارامترها برای یافتن بهترین ساختار، به مدلی دست یافته‌اند که بهترین عملکرد را برای پیش‌بینی قیمت‌های آتی دارد. هرچند که این مطالعه دقت بالای مدل را با استفاده از معیارهای آماری تأیید می‌کند اما برای پیش‌بینی‌های کوتاه‌مدت مناسب هستند و در مواجهه با تغییرات ناگهانی و بزرگ (مانند تغییرات اقتصادی و سیاسی) که خطی نیستند، عملکرد ضعیفی از خود نشان می‌دهند.

در پژوهشی دیگر گریش [۹] بر پیش‌بینی قیمت‌های نقطه‌ای برق در بازار برق هند تمرکز کرده است. این مطالعه با استفاده از داده‌های ساعتی، عملکرد مدل‌های خودرگرسیون GARCH را

[1] Patching and Attention

[2] Volatility

برای این منظور ارزیابی کرده است. با این حال، این مدل‌ها در ثبت کامل پیچیدگی‌ها و الگوهای غیرخطی بازار که ممکن است تحت تأثیر عوامل مختلف قرار گیرد، محدودیت‌هایی دارند و اغلب برای تحلیل واریانس نسبت به پیش‌بینی دقیق قیمت‌ها در بازارهای بسیار پویا مناسب‌تر هستند.

مرابت و همکارانش [۱۰] یک روش ترکیبی جهت بررسی رفتار سری زمانی قیمت نفت، با هدف پیش‌بینی نوسانات آن استفاده کردند. آن‌ها نشان دادند که قیمت نفت، به‌ویژه در بازار مالی، دارای نوسانات بالایی است. آن‌ها با استفاده از داده‌های روزانه از سال ۲۰۱۰ تا ۲۰۱۹، ابتدا سری زمانی را به یک مدل ARIMA تبدیل کرده‌اند. سپس برای مدل‌سازی نوسانات واریانس، از مدل EGARCH استفاده کردند. پس از بررسی و تنظیم پارامترهای مختلف دریافتند که مدل ترکیبی ARIMAو EGARCH بهترین عملکرد را در پیش‌بینی قیمت‌های آتی نفت دارد. اگرچه این مطالعه دقت بالای مدل ترکیبی را تأیید می‌کند، اما اشاره به این نکته که تحقیقات کمی در مورد مدل‌های متقارن و نامتقارن برای نوسانات قیمت نفت وجود دارد، نشان‌دهنده نیاز به بررسی بیشتر در این حوزه است.

رهنمون و همکارانش [۱۱] از یک مدل خودرگرسیون برداری (VAR) برای پیش‌بینی قیمت‌های ماهانه فولاد در ایران استفاده کردند. آن‌ها در مطالعه خود، ضمن شناسایی متغیرهای مؤثر بر قیمت فولاد بر روی داده‌های دوره ۱۳۹۸ تا ۱۴۰۲، نشان دادند که شوک‌های وارد شده از متغیرهای نرخ ارز غیررسمی و شاخص قیمت تولیدکننده صنعت، بیشترین تأثیر را بر نوسانات قیمت فولاد داشته است. آن‌ها با کسب دقت بالای مدل، بر مزیت مدل‌های VAR به دلیل امکان بررسی تأثیر هم‌زمان چند متغیر، نسبت به مدل‌های تک‌متغیره مانند شبکه‌های عصبی یا الگوریتم ژنتیک، تاکید کردند.

در پژوهشی خان و همکارانش [۱۲] عملکرد مدل‌های VAR و ARIMA در پیش‌بینی متغیرهای اقتصادی بنگلادش را مورد مقایسه قرار دادند. هدف اصلی آن‌ها، پیش‌بینی هم‌زمان یک گروه از متغیرها و بهره‌گیری از هم‌بستگی بین آن‌ها بوده است. نتایج این مطالعه نشان داد که مدل VAR در مقایسه با مدل ARIMA، در پیش‌بینی متغیرهایی با هم‌بستگی بالا عملکرد بهتری دارد. با این حال، برای متغیرهایی که هم‌بستگی کمتری با یکدیگر دارند، هر دو مدل VAR و ARIMA عملکرد تقریباً مشابهی داشته‌اند. این یافته‌ها نشان داد که قبل از انتخاب مدل پیش‌بینی، بررسی میزان هم‌بستگی بین متغیرها از اهمیت بالایی برخوردار است، چرا که مدل‌های چندمتغیره مانند VAR برای داده‌های با هم‌بستگی بالا کارایی بیشتری دارند.

۲-۲ پیش‌بینی سری‌های زمانی با روش‌های یادگیری عمیق

پژوهش‌های اخیر نشان داده‌اند که رویکردهای یادگیری ماشین و یادگیری عمیق در مقایسه با مدل‌های سنتی، کارایی بالاتری در پیش‌بینی قیمت‌های غیرخطی و نوسانی دارند. به عنوان مثال، در مقاله [۱۳]، ویجه و همکارانش از دو تکنیک شبکه عصبی مصنوعی و جنگل تصادفی برای پیش‌بینی قیمت‌های پایانی سهام استفاده کردند. این تحقیق با استفاده از داده‌های مالی پنج شرکت، به این نتیجه رسید که مدل شبکه عصبی عملکرد پیش‌بینی بهتری را ارائه می‌دهد. در مقاله‌ای دیگر کوالکار و همکارانش [۱۴] از الگوریتم رگرسیون درخت تصمیم برای پیش‌بینی قیمت مسکن در شهر بمبئی استفاده کردند و توانستند به دقت ۸۹ درصدی دست یابند. این مطالعات تأیید می‌کنند که روش‌های یادگیری ماشین، به ویژه برای داده‌های غیرخطی و پیچیده مالی، می‌توانند ابزارهای قدرتمندی برای پیش‌بینی قیمت‌ها باشند. با در نظر گرفتن این پیش‌زمینه، مطالعات یادگیری عمیق در این حوزه را می‌توان به سه دسته‌بندی تحلیلی اصلی مدل‌های بازگشتی، مدل‌های ترنسفورمر و روش‌های ترکیبی تقسیم کرد. در ادامه برای این دسته‌بندی نمونه‌هایی آورده شده است.

۲-۲-۱ پژوهش‌های مبتنی بر مدل‌های بازگشتی (RNN/LSTM/GRU)

در مقاله [۱۵] شاهی و همکارانش، عملکرد مدل‌های یادگیری عمیق بازگشتی LSTM و GRU برای پیش‌بینی قیمت سهام، تحت شرایط یکسان و کنترل‌شده را مورد بررسی و مقایسه قرار دادند. آن‌ها نشان دادند که هر دو مدل LSTM و GRU، به تنهایی و با استفاده از ویژگی‌های تاریخی سهام، در پیش‌بینی قیمت‌ها

عملکردی وابسته به شرایط دارند. با این حال، هنگامی که داده‌های مربوط به احساسات[1] اخبار مالی به ویژگی‌های ورودی مدل‌ها اضافه می‌شود، عملکرد پیش‌بینی قیمت سهام به طور قابل‌توجهی بهبود می‌یابد. نتایج این تحقیق نشان داد که هر دو مدل ترکیبی LSTM و GRU با داده‌های احساسات توانسته‌اند عملکرد یکسانی را در پیش‌بینی بهتر قیمت سهام ارائه دهند. آن‌ها بر اهمیت گنجاندن داده‌های احساسات در کنار داده‌های تاریخی برای افزایش دقت پیش‌بینی در بازارهای مالی غیرخطی و نوسانی تأکید کردند. گونارتو و همکارانش [۱۶] به مقایسه عملکرد دو مدل یادگیری عمیق RNN و LSTM برای پیش‌بینی قیمت رمزارزها پرداخته‌اند. آن‌ها داده‌های قیمت روزانه بیت‌کوین و اتریوم را از سال ۲۰۱۵ تا ۲۰۲۱ جمع‌آوری و برای پیش‌بینی قیمت‌های آینده استفاده کرده‌اند. نتایج تحقیق نشان می‌دهد که مدل LSTM عملکرد بهتری نسبت به مدل RNN داشته است. این یافته نشان می‌دهد که مدل LSTM ابزار مؤثرتری برای پیش‌بینی قیمت رمزارزها است و می‌تواند به سرمایه‌گذاران در اتخاذ استراتژی‌های مناسب و مدیریت ریسک کمک کند.

۲-۲-۲ پژوهش‌های مبتنی بر مدل‌های ترنسفورمر

مدل ترنسفورمر برای اولین بار برای پیش‌بینی قیمت سهام در بورس داکا بنگلادش به کار گرفته شد [۱۷]. این مدل با استفاده از یک مکانیزم پیشرفته برای نمایش ویژگی‌های سری زمانی، قادر به تحلیل مؤثر داده‌ها است. محققان این مدل را بر روی داده‌های روزانه و هفتگی هشت شرکت منتخب آزمایش کردند و نتایج، عملکرد قابل قبولی را در اکثر سهام‌ها نشان داد. این مطالعه بر شکاف تحقیقاتی موجود در استفاده از مدل‌های مبتنی بر توجه برای پیش‌بینی قیمت سهام در بازار بورس بنگلادش تأکید کرده و با ارائه نتایج امیدوارکننده، پتانسیل بالای این مدل‌ها را در تحلیل و پیش‌بینی سری‌های زمانی پیچیده اثبات می‌کند.

نقطه قوت معماری ترنسفورمر و مکانیزم توجه به خود، امکان پردازش موازی و درک وابستگی‌های بلندمدت بدون محدودیت‌های ذاتی شبکه‌های عصبی بازگشتی RNN است. همین موفقیت باعث شده تا مدل‌های مبتنی بر ترنسفورمر و انواع مختلف آن به حوزه سری‌های زمانی نیز راه یابند. زو و همکارانش [۱۸] با معرفی مدل Informer کارایی این معماری را در پیش‌بینی سری‌های زمانی بلندمدت به اثبات رساندند. در همین راستا، مدل PatchTST [۱۹] به عنوان یک پیاده‌سازی پیشرفته و بهینه از ترنسفورمر برای سری‌های زمانی، با تقسیم داده‌ها به قطعات، توانایی چشم‌گیر مدل در یادگیری الگوهای محلی و بلندمدت را نشان می‌دهد.

۲-۲-۳ پژوهش‌های مبتنی بر روش‌های ترکیبی (مدل‌های کانولوشنی، بازگشتی و ترنسفورمر)

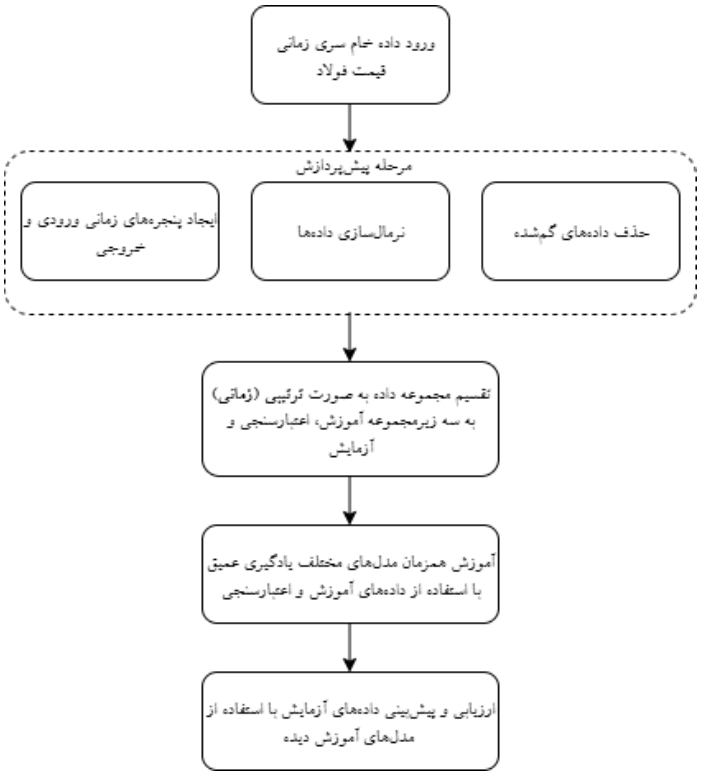
با پیشرفت فناوری، مدل‌های ترکیبی[2] و مکانیزم‌های جدیدی برای بهبود عملکرد مدل‌های یادگیری عمیق در پیش‌بینی سری‌های زمانی معرفی شدند. به عنوان مثال، پاتنایک و همکارانش [۲۰] با استفاده از یک مدل ترکیبی شبکه‌های عصبی کانولوشنال، LSTM و الگوریتم تکاملی گرگ خاکستری، به پیش‌بینی قیمت سهام پرداختند. آن‌ها نشان دادند که ترکیب این معماری‌ها می‌تواند ویژگی‌های مهم‌تری را از داده‌ها استخراج کرده و دقت پیش‌بینی را افزایش دهد. در پژوهشی دیگر، جیالینو همکارانش [۲۱] از ترکیب یک مدل CNN-LSTM با مکانیزم توجه[3] برای پیش‌بینی قیمت سهام استفاده کردند و نشان دادند که مکانیزم توجه می‌تواند به مدل در شناسایی نقاط کلیدی داده‌ها کمک کند و عملکرد آن را بهبود بخشد. هم‌چنین، ونگ [۲۲] یک روش جدید با ترکیب Bi-LSTM و ترنسفورمر برای بهبود دقت و پایداری پیش‌بینی قیمت سهام معرفی کردند. این مدل ترکیبی از Bi-LSTM (برای جذب اطلاعات دوطرفه در توالی‌ها) و یک نسخه بهبودیافته از ترنسفورمر تشکیل شده است. نتایج این مطالعه نشان می‌دهد که این روش جدید، عملکرد بهتری نسبت به روش‌های موجود دارد و با کاهش خطا و افزایش دقت، قابلیت تعمیم بالایی در پیش‌بینی قیمت سهام دارد. این مطالعات تأیید می‌کنند که ادغام معماری‌های مختلف یادگیری عمیق، رویکردی مؤثر برای افزایش دقت در پیش‌بینی سری‌های زمانی پیچیده است.

[1] Sentiment
[2] Hybrid
[3] Attention

با توجه به بررسی مطالعات انجام شده در پیش‌بینی قیمت فولاد، یک خلأ پژوهشی بزرگی مانند فقدان یک مقایسه جامع بین معماری‌های نوین یادگیری عمیق (به ویژه مدل‌های ترنسفورمر) در حوزه تخصصی قیمت فولاد به چشم می‌خورد. لذا، این پژوهش مستقیماً با هدف پر کردن این خلأ، کارآمدترین معماری را برای مدل‌سازی دینامیک‌های پیچیده قیمت فولاد شناسایی می‌کند. هم چنین با توجه به محدودیت‌های روش‌های سنتی در تحلیل الگوهای غیرخطی و پیچیده سری‌های زمانی و با الهام از موفقیت چشم‌گیر مدل‌های یادگیری عمیق در این زمینه، در پژوهش حاضر تلاش شده است تا با استفاده از مدل‌های یادگیری عمیق مانند RNN، LSTM، GRU، Bi-LSTM، TCN ، RCNN، Transformer و PatchTST به پیش‌بینی قیمت فولاد پرداخته شود. از آنجا که هدف این تحقیق ارزیابی عملکرد این معماری‌ها بر روی یک سری زمانی تکی است، این پژوهش به صورت تک‌متغیره انجام شده تا مقایسه‌ای دقیق از کارایی ذاتی هر مدل ارائه گردد و کارآمدترین رویکرد برای پیش‌بینی قیمت فولاد شناسایی شود.

۳- روش پیشنهادی

در این بخش، رویکرد پژوهش برای پیش‌بینی سری‌های زمانی قیمت فولاد که شامل مراحل اصلی آماده‌سازی و پیش‌پردازش مجموعه داده، معرفی معماری مدل‌های یادگیری عمیق مورد استفاده و تشریح معیارهای ارزیابی و تنظیمات پارامترهای آموزش است، با جزئیات کامل شرح داده‌شده است. خلاصه روش پیشنهادی در شکل (۱) آورده شده است.



شکل (۱): خلاصه روش پیشنهادی

۳-۱ آماده‌سازی و پیش‌پردازش مجموعه داده

مجموعه داده مورد استفاده در این پژوهش با عنوان "TATASTEEL.NS_stock_data.csv"[1] از پلتفرم معتبر Kaggle تهیه‌شده است. این داده‌ها شامل اطلاعات تاریخی عملکرد سهام شرکت چندملیتی هندی در صنعت فولاد است. این مجموعه داده به تحلیلگران و پژوهشگران امکان می‌دهد تا روندهای گذشته و نوسانات قیمت سهام را در یک بازه زمانی مشخص بررسی کنند. این اطلاعات می‌تواند برای تحلیل سرمایه‌گذاری، تحقیقات مالی، و پیش‌بینی روندهای بازار به کار گرفته شود. مجموعه داده مذکور، اطلاعات روزانه قیمت سهام را در بازه زمانی ۳ جولای ۲۰۱۵ تا ۲ جولای ۲۰۲۱ ارائه می‌دهد. جدول (۱) ستون‌های کلیدی موجود در مجموعه داده را نشان می‌دهد.

جدول (۱): ستون‌های موجود در مجموعه داده مورد استفاده

مفهوم ستون	نام ستون
تاریخ روزانه	Date
قیمت باز شدن فولاد	Open
بالاترین قیمت در طول روز	High
پایین‌ترین قیمت در طول روز	Low
قیمت بسته شدن فولاد	*Close
حجم معاملات روزانه	Volume

با توجه به جدول (۱) قیمت بسته شدن فولاد به عنوان متغیر اصلی و هدف پیش‌بینی در این پژوهش مورد استفاده قرارگرفته است. مجموعه داده معرفی‌شده به عنوان یک سری زمانی مالی پرنوسان مرتبط با فولاد شناخته‌شده است که دلیل اصلی جهت انتخاب مجموعه داده اصلی برای این پژوهش است. لازم به ذکر است پژوهش جاری جهت تعمیم‌پذیری بر روی دو مجموعه داده مستقل دیگر نیز اجرا و ارزیابی‌شده است، که شامل قیمت آتی کالای جهانی فولاد و یک شاخص کلان ماهانه قیمت فولاد هستند.

پیش‌پردازش داده‌ها به منظور آماده‌سازی آن‌ها برای مدل‌های یادگیری عمیق، در سه مرحله انجام شده است:

۱- حذف داده‌های گم‌شده: ابتدا، تمام ردیف‌هایی که شامل مقادیر گم‌شده هستند، از مجموعه داده حذف شدند تا از یکپارچگی داده‌ها اطمینان حاصل شود.

۲- نرمال‌سازی داده‌ها: برای کاهش حساسیت مدل‌ها نسبت به مقیاس داده‌ها و بهبود همگرایی در طول آموزش، از روش نرمال‌سازی MinMaxScaler استفاده شد. این روش مقادیر قیمت را به بازه [۰،۱] نگاشت می‌کند. نحوه نرمال‌سازی در رابطه (۱) آورده شده است:

$$x'_{MinMax} = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (۱)$$

۳- ایجاد پنجره‌های زمانی: در این مرحله داده‌های سری زمانی به پنجره‌های ورودی و خروجی تبدیل شدند. یک پنجره با طول seq_len برابر با ۳۰ (قیمت ۳۰ روز گذشته) به عنوان ورودی و یک پنجره با طول pred_len برابر با ۷ (قیمت ۷ روز آینده) به عنوان خروجی تعریف شده است. لازم به ذکر است انتخاب این اعداد با استفاده از مفهوم اعتبارسنجی متقابل[2] انجام شده است. درنهایت، این داده‌ها به سه مجموعه آموزش (۷۰٪)، اعتبارسنجی (۱۵٪) و آزمون (۱۵٪) تقسیم شده‌اند.

۳-۲ معرفی معماری مدل‌های یادگیری عمیق مورد استفاده

انتخاب مدل‌های یادگیری عمیق در این پژوهش بر اساس توانایی ساختاری و عملکردی آن‌ها در مواجهه با چالش‌های اساسی سری‌های زمانی مالی پرنوسان قیمت فولاد (یا سهام مرتبط با آن) صورت گرفته است. این چالش‌ها عمدتاً شامل نویز بالا، الگوهای غیرخطی و حضور وابستگی‌های زمانی با طول‌های متفاوت است. برای پوشش‌دهی این طیف از ویژگی‌ها، سه خانواده اصلی مدل‌ها برای مقایسه عملکرد انتخاب شدند: مدل‌های بازگشتی برای ارزیابی و مدل‌سازی وابستگی‌های کوتاه‌مدت و میان‌مدت و تأثیر توالی مستقیم داده‌ها؛ مدل‌های کانولوشنی به دلیل توانایی در

[1] https://www.kaggle.com/datasets/nitirajkulkarni/tata steel-ns-stock-performance

[2] Cross Validation

استخراج ویژگی‌های محلی موضعی (نظیر نوسانات کوچک) و استفاده از Dilated Convolutions برای گسترش مؤثر میدان دید در توالی‌های زمانی؛ و درنهایت، مدل‌های مبتنی بر ترنسفورمر که هدف اصلی پژوهش را تشکیل می‌دهند، برای بررسی کارایی مکانیزم توجه به خود در کشف و مدل‌سازی همبستگی‌های بسیار بلندمدت و غیرخطی داده‌های قیمت فولاد انتخاب گردیدند. به‌طور ویژه، PatchTST با رویکرد پچ‌سازی، باعث مدل‌سازی مؤثرتر الگوهای محلی و بلندمدت در سری‌های زمانی با طول توالی زیاد می‌شود. در این بخش، به بررسی دقیق‌تر معماری مدل‌های یادگیری عمیقی پرداخته‌شده است که برای تحلیل و پیش‌بینی سری‌های زمانی قیمت فولاد مورد استفاده قرارگرفته‌اند. این مدل‌ها به دلیل توانایی در یادگیری الگوهای پیچیده و وابستگی‌های زمانی، از مدل‌های سنتی عملکرد بهتری دارند.

مدل‌های شبکه عصبی بازگشتی:

شبکه‌های عصبی بازگشتی (RNN) دسته‌ای از شبکه‌های عصبی هستند که به دلیل ماهیت ترتیبی خود، برای تحلیل داده‌های سری زمانی مناسب هستند. این مدل‌ها با استفاده از یک حلقه بازگشتی، اطلاعات را از گام‌های زمانی قبلی به گام‌های بعدی منتقل می‌کنند.

الف. RNN ساده: این مدل به عنوان پایه‌ای‌ترین مدل در این دسته، اطلاعات را از گام زمانی قبلی به گام زمانی فعلی منتقل می‌کند. این مدل یک حلقه بازگشتی دارد که به آن اجازه می‌دهد تا اطلاعات را در حافظه خود نگه دارد. RNN در مدل‌سازی وابستگی‌های کوتاه مدت عملکرد خوبی دارد. با این حال، مهم‌ترین ضعف آن، پدیده محو شدن گرادیان است. این پدیده باعث می‌شود که مدل نتواند وابستگی‌های بلندمدت را به خوبی یاد بگیرد، زیرا تأثیر اطلاعات دوردست در طول زمان از بین می‌رود [۵]. به همین دلیل، در مسائل پیچیده سری زمانی، از مدل‌های پیشرفته‌تر استفاده می‌شود.

ب. LSTM: این مدل که توسط Hochreiter و Schmidhuber معرفی شد، یک راه‌حل کارآمد برای مشکل محو شدن گرادیان در RNN است [۶]. این مدل از یک ساختار پیچیده‌تر به نام سلول حافظه استفاده می‌کند که قادر به ذخیره و بازیابی اطلاعات برای مدت‌های طولانی است. LSTM با استفاده از سه دروازه اصلی، جریان اطلاعات را به دقت کنترل می‌کند: دروازه فراموشی[1]، دروازه ورودی[2] و دروازه خروجی[3]، این مکانیزم به LSTM اجازه می‌دهد تا اطلاعات مهم را برای مدت‌های طولانی به خاطر بسپارد و وابستگی‌های پیچیده را به خوبی مدل‌سازی کند.

ج. GRU: این مدل که توسط Cho معرفی شد، نسخه‌ای ساده‌تر و کارآمدتر از LSTM است [۶]. این مدل با ترکیب دو دروازه (دروازه ورودی و فراموشی) به یک دروازه به‌روزرسانی[4]، پیچیدگی محاسباتی کمتری نسبت به LSTM دارد. GRU با داشتن تنها دو دروازه (به‌روزرسانی و ریست[5]) سریع‌تر از LSTM آموزش می‌بیند و پارامترهای کمتری دارد، اما در اکثر موارد، عملکردی مشابه با LSTM ارائه می‌دهد. این ویژگی، GRU را به یک انتخاب محبوب در مواقعی که سرعت و کارایی محاسباتی اهمیت دارد، تبدیل کرده است.

د. LSTM دو طرفه[6]: این مدل یک تغییر ساختاری از LSTM است که توالی داده‌ها را در هر دو جهت (به جلو و عقب) پردازش می‌کند [۲۰]. این مدل از دو لایه LSTM استفاده می‌کند: یکی که توالی را از ابتدا تا انتها می‌خواند و دیگری که آن را از انتها به ابتدا می‌خواند. Bi-LSTM با ترکیب خروجی‌های این دو لایه، به یک درک جامع و کامل از هر گام زمانی دست پیدا می‌کند، زیرا اطلاعات را هم از گذشته و هم از آینده توالی (بخش‌های بعدی) دریافت می‌کند. این ویژگی می‌تواند در مسائلی که وابستگی‌های دوطرفه در داده‌ها وجود دارد، دقت پیش‌بینی را به شکل چشم‌گیری بهبود بخشد.

مدل‌های ترنسفورمر:

ترنسفورمرها، که در ابتدا توسط Vaswani برای پردازش زبان طبیعی توسعه یافتند، با استفاده از مکانیزم توجه به خود، انقلابی در پردازش داده‌های ترتیبی ایجاد کردند. برخلاف RNN که داده‌ها را به صورت متوالی پردازش می‌کند، ترنسفورمر می‌تواند همه نقاط داده را به صورت موازی تحلیل کرده و وابستگی بین آن‌ها را فارغ از فاصله‌شان درک کند. این معماری به دلیل کارایی بالا و توانایی

[1] Forget Gate
[2] Input Gate
[3] Output Gate
[4] Update Gate
[5] Reset Gate
[6] Bi-LSTM

در مدل‌سازی وابستگی‌های بلندمدت، به سرعت به حوزه سری‌های زمانی راه یافت.

مدل‌های ترنسفورمر از دو بخش اصلی رمزگذار[1] و رمزگشا[2] تشکیل شده‌اند. بلوک رمزگذار مسئول پردازش ورودی‌ها و تولید بازنمایی‌های داخلی است. این بلوک شامل لایه‌های توجه چندسر[3] و شبکه‌های پیش‌خور موضعی است. بلوک رمزگشا مسئول تولید خروجی‌ها بر اساس بازنمایی‌های داخلی تولیدشده توسط رمزگذار است [۱۷].

مفهوم توجه چندسر، خود از تعدادی لایه توجه به خود تشکیل شده است. هر یک از این لایه‌ها میزان توجه و ارزش‌گذاری بر روی قسمت‌های مختلف ورودی را با استفاده از سه ماتریس پرسش[4]، کلید[5] و ارزش[6] متفاوت انجام داده و یک بازنمایی منحصر به فردی از ورودی را ایجاد می‌کنند. بازنمایی‌های مختلف به دست آمده از لایه‌های خود توجه با هم ترکیب شده تا بازنمایی قوی‌تر از ویژگی‌ها به دست آید. در ادامه با اضافه کردن لایه‌های پیش‌خور، بازنمایی‌های تولیدشده به خروجی مناسب نگاشت می‌شوند. نحوه عملکرد مفهوم توجه به خود در رابطه (۲) آورده شده است:

$$\begin{aligned} \text{head}_i &= \text{Attention}(QW^Q, KW^K, VW^V) \\ &= \text{Softmax}\left[\frac{QW_i^Q(KW_i^K)^T}{\sqrt{(d_k)}}\right]VW_i^V \\ \text{MultiHead}(Q,K,V) &= \\ \text{concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_o & \end{aligned} \quad (۲)$$

این فرمول نشان می‌دهد که چگونه وابستگی بین ورودی‌ها محاسبه می‌شود. پس از آن، لایه توجه چندسر، چندین بار این عملیات را به صورت موازی انجام می‌دهد و خروجی‌ها را به هم می‌چسباند تا یک نمایش کامل‌تر از ورودی به دست آید.

مدل PatchTST[7]: این مدل یک رویکرد پیشرفته و قدرتمند در حوزه پیش‌بینی سری‌های زمانی است که معماری انقلابی ترنسفورمر را برای این نوع داده‌ها بهینه کرده است. نقطه قوت و نوآوری اصلی این مدل، استفاده از رویکرد قطعات یا پچینگ است. در این روش، به جای اینکه مدل هر نقطه زمانی را به صورت جداگانه پردازش کند، سری زمانی به قطعات کوچکی با طول ثابت تقسیم می‌شود. این پچ‌ها سپس به صورت موازی به مدل ترنسفورمر داده می‌شوند. این کار سه مزیت کلیدی دارد:

۱- استخراج ویژگی‌های محلی: هر پچ به مدل اجازه می‌دهد تا الگوهای محلی (مانند تغییرات کوچک یا روندهای کوتاه) را به صورت مؤثرتر استخراج کند، مشابه کاری که شبکه‌های عصبی کانولوشنال در پردازش تصویر انجام می‌دهند.

۲- کاهش پیچیدگی محاسباتی: با کاهش طول توالی از تعداد کل نقاط داده به تعداد پچ‌ها، پیچیدگی محاسباتی مکانیزم توجه به خود ترنسفورمر به شکل چشم‌گیری کاهش می‌یابد و سرعت آموزش و پیش‌بینی افزایش پیدا می‌کند.

۳- مدل‌سازی وابستگی‌های بلندمدت: مکانیزم توجه به خود به مدل اجازه می‌دهد تا وزن هر پچ را نسبت به پچ‌های دیگر در کل توالی مشخص کند و وابستگی‌های بلندمدت را بدون از دست دادن اطلاعات، کشف نماید.

مدل PatchTST هم‌چنین از یک لایه RevIN برای نرمال‌سازی درونی استفاده می‌کند که پایداری آموزش را بهبود می‌بخشد و حساسیت مدل به تغییرات مقیاس داده‌ها را کاهش می‌دهد. این مدل با ترکیب قدرت ترنسفورمر و رویکرد نوآورانه تقسیم‌بندی به قطعات، عملکردی برتر در پیش‌بینی دقیق سری‌های زمانی، به‌ویژه در مقایسه با مدل‌های قدیمی‌تر، ارائه می‌دهد [۲۰-۲۲].

مدل‌های مبتنی بر شبکه‌های کانولوشنی

شبکه‌های عصبی کانولوشنال برای سری‌های زمانی TCN[8]: این مدل یک معماری است که از شبکه‌های عصبی کانولوشنال برای پردازش داده‌های سری زمانی استفاده می‌کند [۲۳]. ویژگی کلیدی TCN، استفاده از پیچش‌های سببی متسع‌شده[9] است که به آن اجازه می‌دهد تا میدان دید خود را برای دربرگرفتن وابستگی‌های بلندمدت گسترش دهد. این معماری می‌تواند در شرایطی که الگوهای محلی در طول زمان تکرار می‌شوند، مفید باشد. با این

[1] Encoder
[2] Decoder
[3] Multi-Head Attention
[4] Query
[5] Key
[6] Value
[7] Patch-wise Time Series Transformer (PatchTST)
[8] Temporal Convolutional Networks (TCN)
[9] Dilated Causal Convolutions

حال، در مسائل پیچیده‌ای که وابستگی‌های غیرخطی و دوردست غالب هستند، ممکن است به اندازه مدل‌های مبتنی بر توجه مانند ترنسفورمرها عملکرد قوی‌ای ارائه ندهد.

شبکه‌های عصبی بازگشتی-کانولوشنال CRNN[1]:

CRNN یک مدل ترکیبی است که از ترکیب دو معماری، شبکه‌های عصبی کانولوشنال و شبکه‌های عصبی بازگشتی ایجاد شده است [۲۴]. در این معماری، لایه‌های CNN در ابتدا قرار می‌گیرند تا به صورت خودکار ویژگی‌های محلی را از داده‌های ورودی استخراج کنند. پس از این مرحله، خروجی لایه‌های CNN به عنوان ورودی به لایه‌های RNN (مانند LSTM یا GRU) داده می‌شود. لایه‌های RNN مسئول یادگیری وابستگی‌های زمانی و الگوهای ترتیبی از ویژگی‌های استخراج‌شده هستند. این رویکرد دو مرحله‌ای به CRNN اجازه می‌دهد تا هم ویژگی‌های محلی را تشخیص دهد و هم وابستگی‌های زمانی را مدل‌سازی کند. با این حال، در برخی موارد، پیچیدگی این مدل ترکیبی می‌تواند منجر به کاهش کارایی در مقایسه با مدل‌های ساده‌تر و هدفمندتر شود.

۳-۳ معیارهای ارزیابی و تنظیمات پارامترهای آموزش

تمامی مدل‌ها با استفاده از مجموعه‌ای یکسان از تنظیمات آموزش و معیارهای ارزیابی، مورد مقایسه قرار گرفتند تا نتایج به صورت منصفانه و دقیق قابل تحلیل باشند. این رویکرد، امکان مقایسه مستقیم عملکرد مدل‌ها را فراهم می‌کند. جدول (۲) پارامترها و مقادیر آن‌ها را نمایش می‌دهد. توجه به این نکته ضروری است که تنظیمات ساده و یکسان آموزش به طور عمدی انتخاب شدند تا یک مقایسه کنترل‌شده بین توانایی‌های ذاتی معماری مدل‌ها فراهم شده و از هرگونه مزیت ناشی از بهینه‌سازی‌های اختصاصی برای یک مدل خاص جلوگیری شود.

جدول (۲): پارامترهای مدل‌های یادگیری عمیق استفاده‌شده

پارامتر	مقدار
بهینه‌ساز	Adam
نرخ یادگیری	۰/۰۰۰۱
تابع هزینه	MSE
تعداد تکرارهای آموزش	۱۵
اندازه دسته	۳۲

معیارهای ارزیابی:

عملکرد مدل‌ها در مجموعه داده آزمون با سه معیار اصلی ارزیابی‌شده است. این معیارها کمک می‌کنند تا دقت و توانایی مدل در پیش‌بینی قیمت‌های آینده از زوایای مختلف سنجیده شوند.

خطای میانگین مربعات ریشه (RMSE)[2]

این معیار، میانگین اختلاف بین مقادیر واقعی و مقادیر پیش‌بینی‌شده را محاسبه می‌کند. از آنجا که خطاهای بزرگ‌تر را به صورت درجه دوم جریمه می‌کند، نسبت به خطاهای پرت[3] حساسیت بیشتری دارد. هرچه مقدار RMSE کمتر باشد، عملکرد مدل بهتر است. رابطه (۳) نحوه محاسبه این معیار را نمایش داده است.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (۳)$$

خطای میانگین مطلق (MAE)

این معیار، میانگین قدر مطلق اختلاف بین مقادیر واقعی و پیش‌بینی‌شده را نشان می‌دهد. تفسیر MAE ساده‌تر از RMSE است، زیرا هر خطایی را به صورت خطی محاسبه می‌کند . MAE نشان‌دهنده میانگین اندازه خطای پیش‌بینی بدون در نظر گرفتن جهت آن است. رابطه (٤) نحوه محاسبه این معیار را نمایش داده است.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \qquad (٤)$$

[1] Convolutional Recurrent Neural Networks (CRNN)
[2] Root Mean Square Error (RMSE)
[3] Outliers

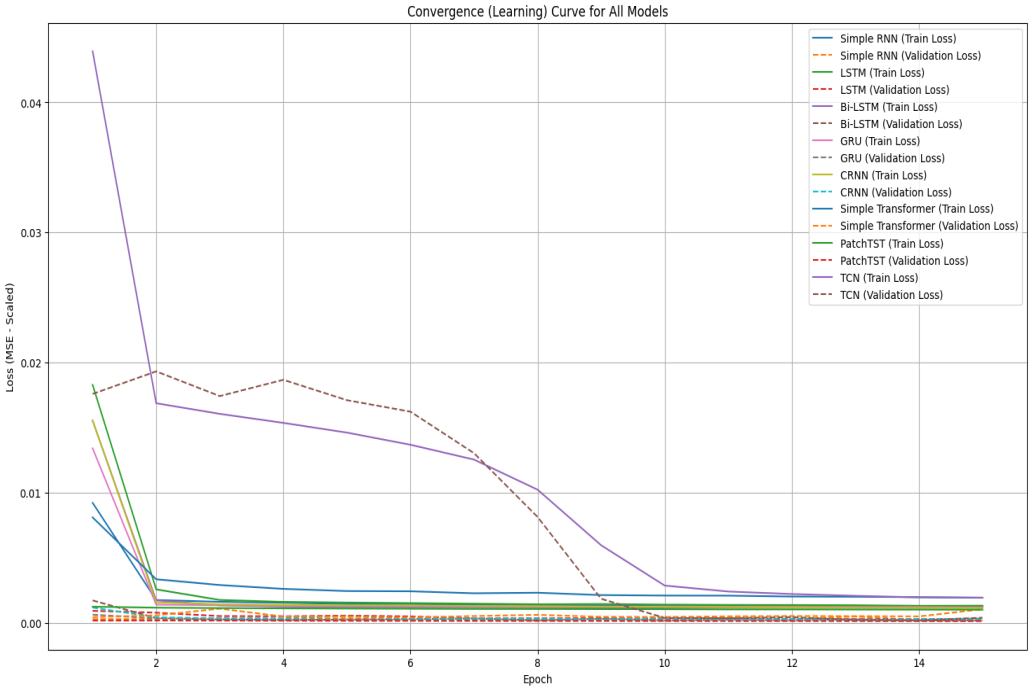
ضریب تعیین (R^2)

این معیار، نشان می‌دهد که چه نسبتی از واریانس داده‌های واقعی توسط مدل قابل توضیح است. R^2 به عنوان یک معیار نرمال‌شده، مقایسه عملکرد مدل‌ها را آسان‌تر می‌کند. مقداری نزدیک به ۱ نشان‌دهنده عملکرد عالی مدل است که تقریباً تمام واریانس داده‌ها را توضیح می‌دهد. مقدار ۰ به این معنی است که مدل هیچ بهبودی نسبت به یک مدل خط پایه ساده (پیش‌بینی میانگین) ندارد، و مقادیر منفی نشان‌دهنده عملکردی بدتر از خط پایه است. هرچه مقدار R^2 به ۱ نزدیک‌تر باشد، مدل از تناسب بهتری با داده‌ها برخوردار است. رابطه (۵) نحوه محاسبه این معیار را نمایش داده است.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \qquad (۵)$$

٤- نتایج ارزیابی

در این پژوهش، برای پیش‌بینی قیمت سهام فولاد، عملکرد چندین مدل یادگیری عمیق ازجمله مدل‌های بازگشتی، مدل‌های مبتنی بر ترنسفورمر و شبکه‌های کانولوشنی مورد بررسی و مقایسه قرارگرفته است. برای حفظ شفافیت، قابلیت تکرارپذیری و تضمین پیش‌بینی واقع‌بینانه، داده‌های سری زمانی پس از پیش‌پردازش شامل حذف داده‌های گم‌شده، نرمال‌سازی با MinMaxScaler و ایجاد پنجره‌های زمانی به صورت کاملاً ترتیبی (زمانی) به سه زیرمجموعه آموزش، اعتبارسنجی و آزمون با نرخ ۱۵/۱۵/۷۰ تقسیم شدند تا از نشت اطلاعات به‌طور کامل جلوگیری شود. تمامی مدل‌ها با استفاده از تنظیمات یکسانی ارزیابی شدند :طول توالی ورودی برابر با ۳۰ روز و طول پیش‌بینی خروجی برابر با ۷ روز در نظر گرفته شد؛ هایپرپارامترهای بهینه‌سازی شامل بهینه‌ساز Adam، نرخ یادگیری 1e-4، اندازه دسته ٦٤ و ۵۰ دوره آموزش بودند. برای اطمینان از ثبات نتایج، هر آزمایش ۳ بار به‌صورت مستقل با مقادیر بذر تصادفی متفاوت تکرار شد. درنهایت، ارزیابی عملکرد مدل‌ها بر روی داده‌های آزمون با استفاده از سه معیار اصلی RMSE، MAE و R^2 انجام شده است. برای ارزیابی همگرایی مدل‌ها و اطمینان از عدم وقوع بیش‌برازش، نمودار منحنی یادگیری تمام مدل‌های یادگیری عمیق درشکل (۲) ترسیم شده است. این نمودار نشان می‌دهد که تابع هزینه فاز آموزش و اعتبارسنجی هم‌زمان کاهش‌یافته و در طول تکرارها تثبیت شده‌اند، که نشان‌دهنده فرایند یادگیری پایدار و بهینه است.



شکل (۲): نمودار همگرایی تابع هزینه فاز آموزش و اعتبارسنجی تمام مدل‌های مورد استفاده

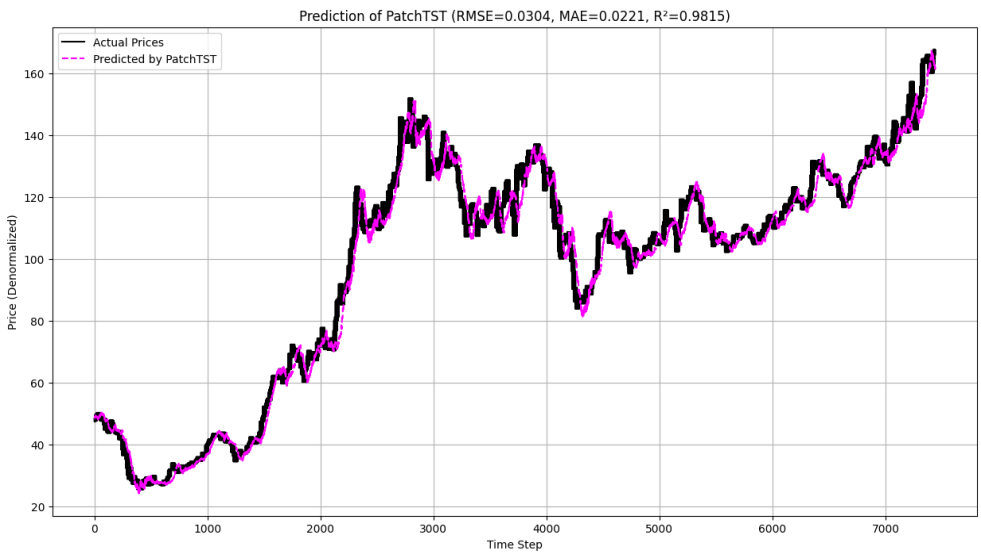
جدول (۳) خلاصه‌ای از میانگین عملکرد حاصل از این ۳ تکرار اجرای مدل‌ها را نشان می‌دهد.

جدول (۳): نتایج ارزیابی میانگین مدل‌های مختلف

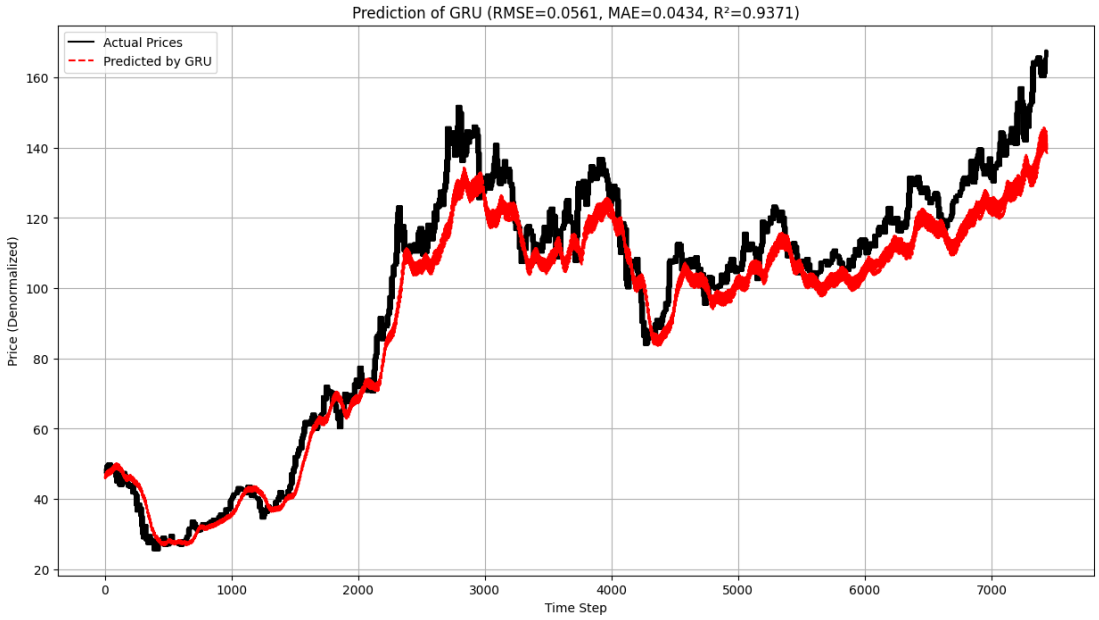
مدل	RMSE	MAE	R^2	زمان آموزش (ثانیه)
CRNN	۰/۱۰۱۲	۰/۰۸۰۱	۰/۷۹۷۵	۱۴/۳۳
TCN	۰/۰۸۸۱	۰/۰۶۸۱	۰/۸۴۵۳	۱۹۶/۴۷
RNN	۰/۱۰۳۴	۰/۰۸۱۴	۰/۷۸۶۸	۱۱/۲۹
LSTM	۰/۰۶۰۴	۰/۰۴۶۹	۰/۹۲۷۲	۱۲/۳۲
Bi-LSTM	۰/۰۸۳۹	۰/۰۶۶۸	۰/۸۵۹۶	۱۵/۸۸
GRU	۰/۰۵۶۱	۰/۰۴۳۴	۰/۹۳۷۱	۱۱/۰۱
Transformer	۰/۰۶۲۸	۰/۰۴۸۰	۰/۹۲۱۲	۳۴/۸۸
PatchTST	۰/۰۳۰۴	۰/۰۲۲۱	۰/۹۸۱۵	۳۷/۱۵

همان‌طور که از جدول (۳) مشخص است، نتایج به وضوح برتری برخی مدل‌ها را نسبت به دیگری نشان می‌دهد. مدل PatchTST با کسب بهترین امتیاز در تمامی معیارها، عملکردی کاملاً برتر را ارائه داد. با RMSE بسیار پایین (۰/۰۳۰۴) و R^2فوق‌العاده بالا (۰/۹۸۱۵)، این مدل توانسته است بیش از ۹۸٪ از تغییرات قیمت فولاد را توضیح دهد. این برتری به دلیل معماری مبتنی بر ترنسفورمر است که با استفاده از مکانیزم توجه به خود، به مدل اجازه می‌دهد تا روابط پیچیده و وابستگی‌های بلندمدت را در سری زمانی کشف کند، بدون اینکه با مشکلات مدل‌های ترتیبی مواجه شود. هم‌چنین، رویکرد تقسیم داده‌ها به قطعات، به این مدل کمک کرده تا الگوهای محلی را با دقت بیشتری یاد بگیرد. در میان مدل‌های بازگشتی، GRUبا RMSE معادل ۰/۰۵۶۱ و R^2برابر ۰,۹۳۷۱، بهترین عملکرد را به نمایش گذاشته است. این نتایج نشان می‌دهد که مدل GRU با وجود سادگی ساختاری، در مقایسه با LSTMاز لحاظ دقت و کارایی محاسباتی، برتری دارد. مدل RNN با ضعیف‌ترین عملکرد R^2 به مقدار ۰/۷۸۶۸ به عنوان یک خط مبنا، به خوبی نشان می‌دهد که مدل‌های پیشرفته‌تر تا چه اندازه توانسته‌اند مشکلات مدل‌های ساده‌تر، به خصوص پدیده محو شدن گرادیان، را حل کرده و دقت پیش‌بینی را به طور چشمگیری افزایش دهند. مدل TCN در بین مدل‌های مبتنی بر شبکه کانولوشن عملکرد متوسطی دارد اما به دلیل استفاده از مفهوم کانولوشن بیشترین زمان آموزش را از بین تمام مدل‌ها به خود اختصاص داده است.

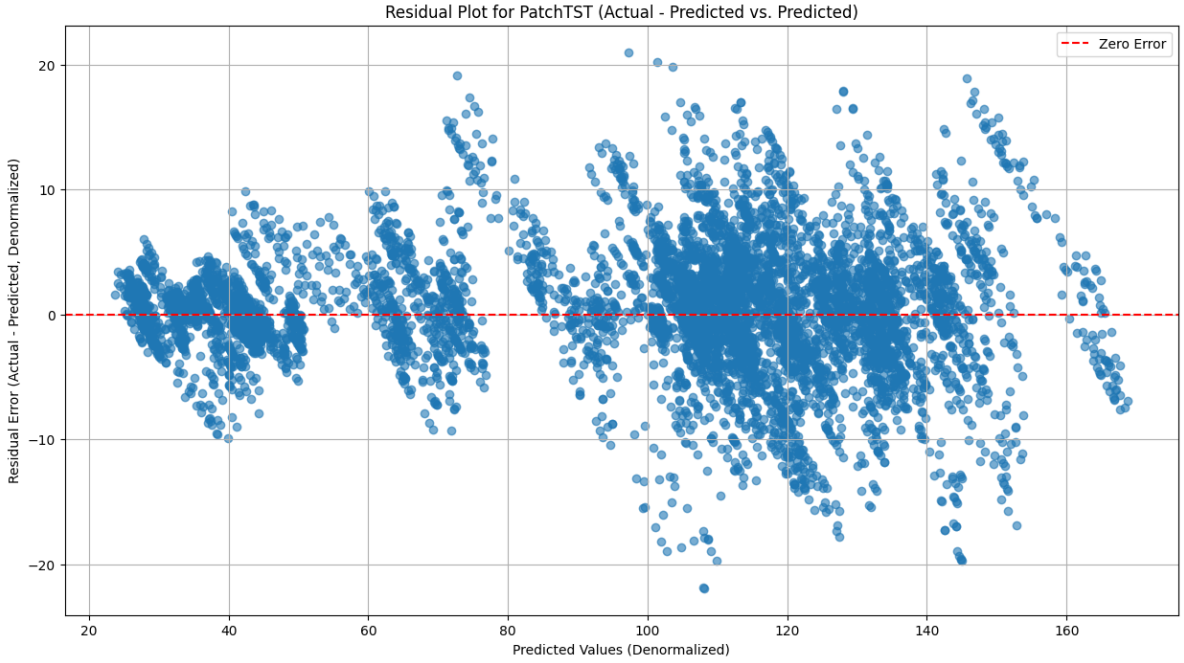
برای درک بهتر عملکرد مدل‌ها، شکل‌های (۳) و (٤) عملکرد دو مدل برتر را ترسیم کرده است.



شکل (۳): مقایسه مقادیر واقعی و پیش‌بینی شده با روش PatchTST



شکل (٤): مقایسه مقادیر واقعی و پیش‌بینی شده با روش **GRU**



شکل (٥): نمودار خطای باقی‌مانده مقدار واقعی و پیش‌بینی شده برای مدل **PatchTST**

این شکل‌ها قیمت‌های پیش‌بینی‌شده را با قیمت‌های واقعی در مجموعه داده آزمون مقایسه کرده است. این نمودارها به صورت بصری نشان می‌دهند که چگونه نمودار پیش‌بینی مدل PatchTST با کمترین انحراف، نزدیک‌ترین تطابق را با داده‌های واقعی داشته است. به منظور بررسی سیستماتیک و سلامت آماری مدل برتر، نمودار خطای باقی‌مانده برای مدل PatchTST درشکل (۵) آورده شده است. توزیع تصادفی خطاها حول خط صفر، شاهدی بر عدم وجود الگوی خطای ساختاریافته و توانایی مدل در ثبت وابستگی‌های غیرخطی بدون سوگیری است.

٤-١ مقایسه روش پیشنهادی با دیگر پایگاه داده‌ها

برای اطمینان از تعمیم‌پذیری و مقاومت روش پیشنهادی، دو پایگاه داده مستقل مرتبط با قیمت واقعی فولاد، علاوه بر داده‌های اصلی سهام فولاد (TATASTEEL.NS_stock_data)، مورد ارزیابی قرار گرفتند. این اقدام به طور خاص جهت پاسخ به دغدغه‌های مربوط به نوع داده و اعتبار علمی پژوهش انجام شده است. مجموعه داده‌های اعتبارسنجی شامل شاخص رسمی قیمت ماهانه ورق فولاد (FRED WPU101707)[1] است که الگوهای کلان را در یک بازه طولانی (از ژوئن ۱۹۸۲ تا می ۲۰۲٤) پوشش می‌دهد و برای این مجموعه داده، پارامترهای بلندمدت ورودی (Seq_len) با مقدار ۱۲ ماه و پارامتر طول خروجی (Pred_len) برابر با ۳ ماه تنظیم شده است. دومین مجموعه داده، قیمت آتی روزانه کلاف نورد گرم (Futures HRC)[2] است که قیمت خام فولاد را با فرکانس روزانه در بازه (از سپتامبر ۲۰۱۵ تا اکتبر ۲۰۲۵) بازتاب می‌دهد و با پارامترهای طول ورودی ۳۰ روزه و طول خروجی ۷ روزه مدل‌سازی شده است. تمام مدل‌ها بر روی هر سه مجموعه داده با ۱۵ تکرار آموزش داده شدند تا مقایسه عملکرد کاملاً منصفانه و مستند باشد. جدول (٤) نتایج مقایسه‌ای جامع مدل‌ها را بر اساس سه معیار اصلی RMSE، MAE و R^2 در سه مجموعه داده با فرکانس‌های را گزارش می‌دهد.

نتایج نهایی موجود در جدول (٤) نشان می‌دهد که در هر سه مجموعه داده متفاوت، مدل PatchTST به‌طور قاطع بهترین عملکرد را ثبت کرده است .این برتری، ریشه در معماری PatchTST دارد که با تقسیم سری زمانی به قطعات و استفاده از مکانیزم توجه، می‌تواند همبستگی‌های بلندمدت و دوردست را به صورت موازی و مؤثر، بدون مشکل محوشدگی گرادیان که محدودیت مدل‌های RNN5/LSTM است، استخراج و پردازش کند. این مقایسه قاطع، فرضیه اصلی پژوهش مبنی بر برتری مدل‌های مبتنی بر ترنسفورمر را بر مدل‌های RNN در پیش‌بینی سری‌های زمانی قیمت فولاد، به دلیل توانایی بالاتر در پردازش موازی و مدیریت وابستگی‌های دوردست، به‌طور کامل اثبات می‌کند.

۵- نتیجه‌گیری

پیش‌بینی دقیق قیمت فولاد به دلیل اهمیت استراتژیک این ماده در صنایع گوناگون، از اهمیت حیاتی برخوردار است. این پژوهش با هدف ارزیابی و مقایسه عملکرد مدل‌های پیشرفته یادگیری عمیق در پیش‌بینی قیمت سری زمانی فولاد انجام شد. در این تحقیق، مجموعه‌ای از مدل‌های بازگشتی، ترنسفورمر و کانولوشنی بر روی داده‌های تاریخی قیمت سهام فولاد مورد آزمایش قرار گرفتند و نتایج به وضوح نشان داد که مدل PatchTST با کسب بهترین امتیاز در تمام معیارهای ارزیابی عملکردی فراتر از سایر مدل‌ها ارائه داده است. این برتری قابل‌توجه، بیش از بازگویی نتایج عددی است و دلیل تحلیلی عمیقی دارد. رویکرد نوآورانه PatchTST در تقسیم سری زمانی به قطعات و استفاده از مکانیزم توجه به خود، به‌طور مؤثر چالش‌های مدل‌های بازگشتی (مانند محوشدگی گرادیان) را حل کرده است؛ این مکانیزم به مدل اجازه می‌دهد که وابستگی‌های پیچیده و بلندمدت را با دقت بالا مدل‌سازی کند، قابلیتی که مدل‌های RNN به دلیل ماهیت ترتیبی و محدودیت‌های ساختاری خود فاقد آن بودند. با وجود این نتایج موفقیت‌آمیز، این پژوهش با چالش تک‌متغیره بودن داده‌ها روبرو است که یک

[1] https://fred.stlouisfed.org/series/WPU101707
[2] https://www.investing.com/commodities/us-steel-coil-futures-historical-data

محدودیت تحلیلی محسوب می‌شود، چرا که عوامل کلان اقتصادی، عرضه و تقاضا، اخبار جهانی و سیاست‌های دولت نیز تأثیر بسزایی در نوسانات قیمت فولاد دارند. از این رو، برای کارهای آینده پیشنهاد می‌شود که از داده‌های چندمتغیره شامل حجم معاملات و شاخص‌های اقتصادی مرتبط استفاده شود، همچنین بررسی امکان استفاده از تکنیک‌های پردازش زبان طبیعی برای تحلیل اخبار و گزارش‌های مالی و ترکیب آن‌ها با داده‌های سری زمانی می‌تواند دقت پیش‌بینی را بهبود بخشد. هم چنین، انجام یک فرآیند بهینه‌سازی سیستماتیک برای یافتن بهترین هایپرپارامترهای هر مدل و آزمایش با طول‌های ورودی و خروجی متفاوت نیز پیشنهاد می‌گردد.

جدول (٤): مقایسه روش پیشنهادی با استفاده از سه مجموعه داده مختلف

معیار ارزیابی	مدل	TATA STEEL	HRC Futures	FRED WPU101707
R^2	PatchTST	۰/۹۸۱۵	۰/۸۸٦۸	۰/۷۵۲۹
	GRU	۰/۹۳۷۱	۰/۸۵٦٤	−۱/٤۸۳۵
	LSTM	۰/۹۲۷۲	۰/٦۱۵۷	−۲/٤۷۵۰
	Bi-LSTM	۰/۷۸٦۸	۰/۸۵۹٦	−۲/۰٤۷۷
	Simple Transformer	۰/۹۲۱۲	۰/۷۷۵۵	۰/۱۰۸۸
	RNN	۰/۸۵۹٦	۰/۷۵۹۱	۰/۳۲۱۵
	TCN	۰/۸٤۵۳	−۱/٤۹۰۲	−۲/٦۵۳٤
	CRNN	۰/۷٦۷۵	۰/۷۹۷٤	−۰/٦۵۹۱
RMSE	PatchTST	۰/۰۳۰٤	۰/۰۲۰۳	۰/۱۱۳۱
	GRU	۰/۰۵٦۱	۰/۰۲۲۸	۰/۳۵۸۵
	LSTM	۰/۰٦۰٤	۰/۰۳۷٤	۰/٤۲٤۱
	Bi-LSTM	۰/۱۰۳٤	۰/۰۲۲٦	۰/۳۹۷۱
	Simple Transformer	۰/۰٦۲۸	۰/۰۲۸۵	۰/۲۱٤۸
	RNN	۰/۰۸۳۹	۰/۰۲۹٦	۰/۱۸۷٤
	TCN	۰/۰۸۸۱	۰/۰۹۵۱	۰/٤۳٤۸
	CRNN	۰/۱۰۷۹	۰/۰۲۷۱	۰/۲۹۳۰
MAE	PatchTST	۰/۰۲۲۱	۰/۰۱٤۳	۰/۰۷٤۲
	GRU	۰/۰٤۳٤	۰/۰۱۵۰	۰/۲۹٦۱
	LSTM	۰/۰٤٦۹	۰/۰۲٦۰	۰/۳٦۱۲
	Bi-LSTM	۰/۰۸۱٤	۰/۰۱۵٦	۰/۳۳۷۱
	Simple Transformer	۰/۰٤۸۰	۰/۰۱۹۳	۰/۱۳۳۳
	RNN	۰/۰٦٦۸	۰/۰۲٤۸	۰/۱۱۵۱
	TCN	۰/۰٦۸۱	۰/۰۷۹٦	۰/۳۷۲۱
	CRNN	۰/۰۸۸۱	۰/۰۱۷۷	۰/۲۲۱٤

References

[1] G. Dooley, and H. Lenihan, “An assessment of time series methods in metal price forecasting,” Resources Policy, Vol. 30, No. 3, pp. 208-217, 2005.

[2] T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” Journal of econometrics, Vol. 31, No. 3, pp. 307-327, 1986.

[3] C. A. Sims, "Macroeconomics and reality," Econometrica: journal of the Econometric Society, pp. 1-48, 1980.
[4] [4] C. Zhang, N. N. A. Sjarif, and R. Ibrahim, "Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 14, No. 1, 2024.
[5] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation,". arXiv preprint arXiv:1406.1078, 2014.
[6] C. Ozdemir, K. Buluş, and K. Zor, "Medium-to long-term nickel price forecasting using LSTM and networks," Resources Policy, Vol. 78, p. 102906, 2022.
[7] L'Heureux, K. Grolinger, and M. A. Capretz, "Transformer-based model for electrical load forecasting," Energies, Vol. 15, No. 14, p. 4993, 2022.
[8] Guha, and G. Bandyopadhyay, "Gold price forecasting using ARIMA model," Journal of Advanced Management Science, Vol. 4, No. 2, 2016.
[9] G. P. Girish, "Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models," Energy Strategy Reviews, Vol. 11, pp. 52-57, 2016.
[10] F. Merabet, H. Zeghdoudi, R. H. Yahia, and I. Saba, "Modelling of oil price volatility using ARIMA-GARCH models," Adv. Mathematics, Vol. 10, pp. 2361-2380, 2021.
[11] T. Rahnemoon Piruj, S. S. Akbar Mousavi, and M. Asgari, "Modeling and Monthly Price Forecasting of Steel in Iran," Stable Economy Journal, Vol. 4, No. 4, pp. 60-95, 2024.
[12] M. S. Khan, and U. Khan, "Comparison of forecasting performance with VAR vs. ARIMA models using economic variables of Bangladesh," Asian Journal of Probability and Statistics, Vol. 10, No. 2, pp. 33-47, 2020.
[13] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock closing price prediction using machine learning techniques," Procedia computer science, Vol. 167, pp. 599-606, 2020.
[14] Kuvalekar, S. Manchewar, S. Mahadik, and S. Jawale, "House price forecasting using machine learning," In Proceedings of the 3rd international conference on advances in science & technology (ICAST), 2020.
[15] T. B. Shahi, A. Shrestha, A. Neupane, and W. Guo, "Stock price forecasting with deep learning: A comparative study," Mathematics, Vol. 8, No. 9, p. 1441, 2020.
[16] M. Gunarto, S. Sa'adah, and D. Q. Utama, "Predicting cryptocurrency price using rnn and lstm method," Journal Sisfokom (Sistem Informasi dan Komputer), Vol. 12, No. 1, pp. 1-8, 2023.
[17] T. Muhammad, A. B. Aftab, M. Ibrahim, M. M. Ahsan, M. M. Muhu, S. I. Khan, and M. S. Alam, "Transformer-based deep learning model for stock price prediction: A case study on Bangladesh stock market," International Journal of Computational Intelligence and Applications, Vol. 22, No. 3, p. 2350013, 2023.
[18] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting,". In Proceedings of the AAAI conference on artificial intelligence, Vol. 35, No. 12, pp. 11106-11115, 2021.
[19] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, "Transformers in time series: A survey," arXiv preprint arXiv:2202.07125, 2022.
[20] Patnaik, N. J. Rao, B. Padhiari, and S. Patnaik, "Optimised hybrid CNN-LSTM model for stock price prediction,". International Journal of Management and Decision Making, Vol. 23, No. 4, pp. 438-460, 2024.
[21] L. Jialin, Q. Shanwen, Z. Zhikai, L. Keyao, M. Jiayong, and T. T. Toe, "Cnn-lstm model stock forecasting based on an integrated attention mechanism," In 2022 3rd International conference on pattern recognition and machine learning (PRML), pp. 403-408, 2022.
[22] S. Wang, "A stock price prediction method based on BiLSTM and improved transformer,". IEEE Access, Vol. 11, pp. 104211-104223, 2023.
[23] Y. Liu, H. Dong, X. Wang, and S. Han, "Time series prediction based on temporal convolutional network," In 2019 IEEE/ACIS 18th International conference on computer and information science (ICIS), pp. 300-305, 2019.
[24] J. Shi, R. Myana, V. Stebliankin, A. Shirali, and G. Narasimhan, "Explainable parallel rcnn with novel feature representation for time series forecasting," In International Workshop on Advanced Analytics and Learning on Temporal Data, pp. 56-75, 2023.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp.17 -33

Steel Price Time Series Forecasting Using a Patch-based Transformer Architecture

Fatemeh Zare Mehrjardi[1*], Abbas Norouzi[2]

[1]Assistant Professor Computer Engineering Department, Meybod University, Meybod, Iran

[2] M.Sc, Computer Engineering Department, Meybod University, Meybod, Iran

Article Information

Original Research Paper

Received:
2025 August 17

Accepted:
2025 November 04

Keywords:
Steel price forecasting, deep learning, time series, recurrent neural networks, transformer network

Corresponding Author*:
fzare@meybod.ac.ir

Abstract

Steel, as one of the most critical materials across various industries, plays a vital role in the global economy. However, its high price volatility, complex and non-linear dependencies, and the inability of traditional models to capture long-term correlations have made its accurate forecasting a serious challenge. To overcome these shortcomings, this research proposes a comprehensive deep learning approach for steel price time series forecasting, structured in three main phases. First, the data undergoes preprocessing, including handling missing values, normalization, and creating time window inputs. Second, various deep learning models, including Recurrent Neural Networks (RNNs), Transformer networks (PatchTST), and Convolutional Networks, are trained and optimized. Finally, the third phase involves evaluating and comparing the models' performance using the R^2, RMSE, and MAE metrics. Evaluation results demonstrate that the PatchTST model, by utilizing its attention mechanism and processing data in segments or "patches," successfully identified complex, long-term dependencies with superior accuracy compared to other models. Specifically, PatchTST achieved a leading R2 value of 0.9815 and the lowest RMSE of 0.0304 in steel price prediction. Conversely, standard RNN models exhibited significantly weaker performance due to their sequential nature and inherent structural limitations. These findings underscore the definitive superiority of Transformer-based models for accurate forecasting of complex time series data.

: 10.22034/ABMIR.2025.23536.1152

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات۵۲ – ۳۵

تشخیص دو پوسته بودن فولاد با شبکه‌های عصبی مصنوعی: رویکردی مبتنی بر بینایی ماشین و یادگیری عمیق

مصطفی مجیدی[۱*]، سیدامیر مکی‌نژادسانیج [۲]

[۱] شرکت توسعه فولاد آلیاژی ایرانیان، یزد، ایران

[۲] واحد فناوری اطلاعات، شرکت فولاد آلیاژی ایران، یزد، ایران

مقاله پژوهشی

چکیده

در این مقاله رویکردی نوین برای تشخیص خودکار عیب دوپوستگی در میلگردهای فولادی با استفاده از شبکه‌های عصبی مصنوعی و تکنیک‌های بینایی ماشین ارائه‌شده است. عیب دوپوستگی که به‌دلیل جدا شدن لایه‌ای نازک از سطح فولاد در فرآیند نورد ایجاد می‌شود، می‌تواند منجر به کاهش خواص مکانیکی و شکست محصول نهایی گردد. روش‌های سنتی بازرسی که اغلب به‌صورت دستی و زمان‌بر هستند با چالش‌هایی مانند خطای انسانی و سرعت پایین مواجه‌اند. در این پژوهش سامانه‌ای مبتنی بر بینایی ماشین پس از فرایند نورد، جهت تشخیص دوپوستگی قرار داده‌شده است. این سامانه شامل سه دوربین صنعتی با زاویه ۱۲۰ درجه و نورپردازی کنترل‌شده است که امکان تصویربرداری پیوسته و همگام با حرکت خط تولید را فراهم می‌کند. در مرحله اول از تکنیک‌های پیش‌پردازش تصویر نظیر نرمال‌سازی و قطعه‌بندی برای بهبود کیفیت تصویر استفاده‌شده است. سپس مدل‌های یادگیری عمیق، برای طبقه‌بندی و شناسایی عیب دوپوستگی آموزش داده‌شده‌اند. نتایج ارزیابی مدل‌ها بر روی داده‌های واقعی خط تولید نشان داد که مدل **MobileNet** با دستیابی به دقت ۹۱/۲۵ درصد، بهترین عملکرد را در میان مدل‌های مورد بررسی داشته است. نتایج به دست آمده بیانگر این است که سامانه پیشنهادی از قابلیت استقرار صنعتی برخوردار بوده و می‌تواند به‌عنوان جایگزینی مؤثر برای روش‌های سنتی بازرسی عمل کند.

تاریخ دریافت:
۱۴۰۴/۰۵/۳۱
تاریخ پذیرش:
۱۴۰۴/۰۸/۲۲

کلیدواژه‌ها:
تشخیص عیب فولاد، دوپوستگی، شبکه‌های‌عصبی‌مصنوعی، یادگیری عمیق، بینایی ماشین

نویسنده‌مسئول:
m.majidi@stu.yazd.ac.ir

: 10.22034/ABMIR.2025.23565.1159

E-ISSN: 2821-2037

۱- مقدمه

فولاد یکی از رایج‌ترین مواد فلزی در زندگی روزمره ما است که کاربردهای فراوانی دارد و ماده‌ای ایده‌آل در بسیاری از زمینه‌ها است. فولاد به‌طور گسترده در زیرساخت‌های مهندسی عمران، هوافضا، کشتی‌سازی، خودروسازی، تولید ماشین‌آلات و ساخت ابزارهای مختلف خانگی مورد استفاده قرار می‌گیرد. طبق آمارهای انجمن جهانی فولاد، گزارش‌شده است که تقاضای جهانی فولاد در سال‌های اخیر با افزایش چشم‌گیری روبه رو شده است و تولید سالانه آن به ۱۸۴۰ میلیارد تن برسد که بیشتر از مجموع تمام فلزات دیگر است [۱]. بنابراین صنعت فولاد از ارکان بنیادین اقتصاد جهانی است و کیفیت محصولات آن به‌ویژه کیفیت سطح، نقشی تعیین‌کننده در عملکرد، ایمنی و طول عمر سازه‌ها و قطعات ایفا می‌کند. وجود عیوب سطحی مانند ترک‌ها، حفره‌ها، ناخالصی‌ها و به‌ویژه دوپوستگی می‌تواند به کاهش استحکام، افزایش احتمال خستگی و در نهایت شکست محصول منجر شود[۱].

در رویه‌های سنتی کارخانه‌های فولاد، بازرسی سطح به‌صورت بصری و توسط اپراتور انسانی انجام می‌شود. این رویکرد، علاوه بر وابستگی به مهارت و قضاوت‌های فردی با محدودیت‌هایی چون خستگی اپراتور، ناهمسانی تصمیم‌گیری، سرعت پایین نسبت به سرعت خط تولید و پوشش ناقص (نمونه‌برداری تصادفی به‌جای بازرسی صددرصد) مواجه است؛ در نتیجه، احتمال عدم‌تشخیص عیوب کوچک یا پنهان افزایش می‌یابد و میزان دقت آن پایین است [۲].

با شتاب‌گیری تحول دیجیتال و حرکت به‌سوی خودکارسازی فرآیندها، بهره‌گیری از بینایی ماشین و یادگیری عمیق به‌عنوان راهکاری کارآمد برای کنترل کیفیت در خط تولید به‌سرعت موردتوجه قرار گرفت. به عنوان نمونه شبکه‌های عصبی پیچشی[1] در سال‌های اخیر توانسته‌اند در مسائل طبقه‌بندی و آشکارسازی عیوب سطحی عملکردی فراتر از روش‌های کلاسیک ارائه دهند و امکان پیاده‌سازی سامانه‌های بازرسی دقیق و مقیاس‌پذیر را فراهم کنند. چنین سامانه‌هایی علاوه بر حذف خطاهای انسانی ناشی از خستگی، با سرعت خط تولید همگام می‌شوند و پوشش بازرسی ۱۰۰ درصد را ممکن می‌سازند [۳].

در این پژوهش تمرکز بر تشخیص خودکار عیب دوپوستگی در محصولات گرد فولادی (مانند بیلت‌ها و میلگردها) در حین تولید است. ضرورت صنعتی این کار از دو جهت، کلیدی محسوب می‌شود: نخست، تضمین کیفیت محصول نهایی برای مشتری از طریق شناسایی و جداسازی نمونه‌های معیوب پیش از خروج از کارخانه؛ دوم، پیشگیری از آسیب‌های پرهزینه به تجهیزات ایستگاه‌های بعدی در خط تولید، که می‌تواند در اثر برخورد یا درگیری ناحیه دوپوستگی با هد دستگاه‌های آزمون رخ دهد و توقف تولید را به‌دنبال داشته باشد. تجربه عملیاتی نشان می‌دهد که دوپوستگی ایجادشده در مرحله نورد، به‌ویژه در اثر حضور پوسته‌های سطحی فولاد هنگام گرم‌کاری ممکن است در مرحله صافکاری به‌واسطه تنش‌های اعمالی باز شده و به‌صورت پوسته جداشده ظاهر شود؛ همین پدیده هم خطر رد کیفی محصول را بالا می‌برد و هم احتمال آسیب به تجهیزات آزمون و توقف خط تولید را افزایش می‌دهد.

برای پاسخ به این نیاز بعد از واحد عملیات حرارتی و تکمیل‌کاری، سامانه‌ای مبتنی بر بینایی ماشین پس از دستگاه صافکاری پیاده‌سازی شد. پیکربندی سخت‌افزاری شامل سه دوربین صنعتی با آرایش ۱۲۰ درجه به‌منظور پوشش کامل محیطی سطح میلگرد و زیرسامانه نورپردازی کنترل‌شده برای پایدارسازی کیفیت تصویر در برابر تغییرات دمایی و نوری محیط است. این چیدمان امکان تصویربرداری پیوسته و همگام با حرکت خط را فراهم می‌کند تا دوپوستگی به‌صورت زمان‌حقیقی[2] درصد شود و در صورت شناسایی، میلگرد بلافاصله از جریان تولید خارج گردد.

در سطح نرم‌افزاری، چارچوب پیشنهادی بر یک زنجیره کامل داده‌محور استوار است که شامل جمع‌آوری داده در خط تولید، پیش‌پردازش داده (نرمال‌سازی و قطعه‌بندی نواحی موردنظر)، آموزش و تنظیم دقیق مدل‌های یادگیری عمیق برای شناسایی دوپوستگی و ارزیابی مدل است. بدین‌ترتیب علاوه‌بر دستیابی به

[1] Convolutional Neural Networks (CNNs)

[2] Real-Time

دقت بالا در شناسایی، الزامات عملیاتی نظیر تأخیر اندک، پایداری در شرایط واقعی خط و قابلیت استقرار صنعتی نیز برآورده می‌شود. نوآوری‌های اصلی این مقاله را می‌توان به‌صورت زیر خلاصه کرد:

- ارائه یک معماری عملیاتی انتها به انتها[1] برای تشخیص دوپوستگی در محصولات گرد فولادی با پوشش ۳۶۰ درجه سطح میلگرد.
- طراحی خط لوله پیش‌پردازش سازگار با شرایط واقعی خط تولید.
- آموزش و ارزیابی مدل‌های یادگیری عمیق با تکیه بر داده‌های برچسب‌خورده صنعتی.

ساختار این مقاله به شرح زیر است: در بخش ۲، مروری بر پژوهش‌های پیشین ارائه می‌شود. در بخش ۳، مفاهیم پایه مورد نیاز تشریح می‌گردد. بخش ٤ به معرفی روش پیشنهادی اختصاص دارد که شامل فرآیند جمع‌آوری تصاویر، مشخصات دوربین‌ها، پیش‌پردازش داده‌ها و معماری شبکه‌های عصبی مورد استفاده است. در بخش ۵، فرآیند آموزش، داده‌افزایی و معیارهای ارزیابی مورد بحث قرار می‌گیرد. در نهایت، بخش ٦ به جمع‌بندی نتایج و ارائه مسیرهای پژوهشی آینده اختصاص دارد.

۲- مروری بر پژوهش‌های پیشین

تشخیص خودکار عیوب سطح فولاد یکی از بخش‌های حیاتی در کنترل کیفیت محصولات فولادی است. طی دهه‌های اخیر، این حوزه از بازرسی دستی به سمت الگوریتم‌های یادگیری ماشین سنتی و در نهایت یادگیری عمیق پیش رفته است. تشخیص عیوب سطح فولاد، عمدتاً شامل سه نوع وظیفه است: تشخیص، طبقه‌بندی و پیدا کردن محل عیوب[٤]. تشخیص عیوب به معنای تعیین این است که آیا شیء مورد نظر حاوی عیب است یا خیر. طبقه‌بندی عیوب، به معنای اختصاص یک تصویر به یک کلاس معیوب است. و پیدا کردن محل عیوب به معنای تعیین دقیق محل عیوب در تصویر است. در اینجا ما به دنبال تشخیص عیب دوپوستگی فولاد هستیم که در دسته اول روش‌های مزبور قرار می‌گیرد.

روش‌های تشخیص خرابی سطح فولاد به سه دسته روش‌های بازرسی دستی، روش‌های مبتنی بر یادگیری ماشین و روش‌های مبتنی بر یادگیری عمیق تقسیم می‌شوند [۳]. همان‌گونه که بیان گردید دسته اول روش‌های بازرسی دستی بودند که در این روش‌ها، اپراتور انسانی با تکیه بر تجربه و مشاهده بصری، عیوب سطحی را شناسایی می‌کرد. هرچند که این شیوه‌ها در صنایع قدیمی مؤثر بودند، اما محدودیت‌هایی مانند سرعت پایین، خستگی انسانی، و دقت ناکافی از مشکلات آن‌ها است که با افزایش سرعت خطوط تولید این روش‌ها عملاً ناکارآمد شده و راه را برای اتوماسیون و روش‌های خودکار باز کرده است[۳].

دسته دوم، الگوریتم‌های مبتنی بر یادگیری ماشین بودند که از دو مرحله اصلی، استخراج ویژگی و دسته‌بندی تشکیل می‌شوند. اتکا به استخراج ویژگی‌های دستی سبب می‌شود که، این مدل‌ها توانایی تعمیم پایینی داشته باشند و مستعد تداخل و تأثیر نویز شوند، بنابراین دقت تشخیص کاهش می‌یابد.

الگوریتم‌های یادگیری ماشین مورد استفاده برای تشخیص نقص سطح فولاد را می‌توان به‌طور کلی به روش‌های مبتنی بر ویژگی بافت، روش‌های مبتنی بر ویژگی شکل و روش‌های مبتنی بر ویژگی رنگ طبقه‌بندی کرد [۵]. با این حال، در زمینه تشخیص نقص سطح فولاد، از آنجایی که تصاویر صنعتی اغلب به صورت مقیاس خاکستری هستند. لذا روش‌های مبتنی بر ویژگی رنگ در این حوزه کمتر کاربرد دارند.

روش‌های آماری، زیرمجموعه‌ای از روش‌های مبتنی بر ویژگی‌های بافت محسوب می‌شوند که با مدل‌سازی توزیع شدت سطوح خاکستری و روابط آماری بین پیکسل‌های مجاور، ویژگی‌های توصیفی بافت را استخراج می‌کنند. ازجمله شناخته‌شده‌ترین این روش‌ها می‌توان به الگوی دودویی محلی[2] و ماتریس هم‌وقوع سطوح خاکستر[3] اشاره کرد. ویژگی‌های حاصل از این روش‌ها معمولاً به‌عنوان ورودی به الگوریتم‌های یادگیری ماشین مانند SVM یا KNN مورد استفاده قرار می‌گیرند[٦-۸]. برخی از روش‌های دیگر بافت از تحلیل سیگنال تصویر در حوزه فرکانس با فیلترهایی مانند Gabor و Wavelet استفاده می‌کنند.

[1] End-to-End
[2] Local Binary Pattern (LBP)
[3] Gray-Level Co-occurrence Matrix (GLCM)

هدف آن‌ها این است که با تصویر به عنوان یک سیگنال دو بعدی رفتار کنند و سپس تصویر را از نقطه نظر طراحی فیلتر سیگنال تجزیه‌وتحلیل کنند. این روش‌ها از لحاظ پیچیدگی ابعاد بالایی دارند و می‌توانند سربار محاسباتی بالایی داشته باشند[۹-۱۰]. نوع دیگر روش‌های مبتنی بر بافت، روش‌های ساختاری هستند که با تکیه بر الگوهای هندسی مانند لبه‌ها و اسکلت‌بندی به دنبال استخراج ویژگی‌های سطح و خرابی‌های سطحی هستند. این روش‌ها به دلیل وابستگی به ابرپارامترها بسیار حساس هستند و برای هر تصویر باید مقدار آستانه آن‌ها به صورت جدا محاسبه گردد یا از یک الگوریتم بهینه سازی جدا برای پیدا کردن حد آستانه بهینه استفاده گردد که سبب می‌شود هم از نظر زمانی و هم از نظر محاسباتی غیربهینه باشند[۱۱].

دسته دیگری از روش‌های مبتنی بر یادگیری ماشین، روش‌های مبتنی بر شکل هستند. در این روش‌ها، ویژگی‌های تصویری از طریق توصیف‌گرهای شکل[1] استخراج می‌شوند؛ از این‌رو، دقت و کارایی الگوریتم تشخیص نقص به دقت توصیف شکل وابسته است. یک توصیف‌گر شکل مناسب باید دارای ویژگی‌هایی همچون تغییرناپذیری هندسی، انعطاف‌پذیری، انتزاع‌پذیری، یکتایی و جامعیت باشد. به‌طور کلی، توصیف‌گرهای شکل را می‌توان به دو دسته تقسیم کرد، نخست، توصیف‌گرهای کانتور[2] که به توصیف لبه یا مرز بیرونی ناحیه شیء می‌پردازند و دوم، توصیف‌گرهای ناحیه‌ای[3] که کل ناحیه شیء را در فرایند توصیف در نظر می‌گیرند. با وجود کارایی این روش‌ها، وابستگی زیاد آن‌ها به نوع و پارامترهای توصیف‌گر سبب می‌شود که برای هر کاربرد و حتی هر تصویر، نیاز به تنظیم دقیق وجود داشته باشد؛ امری که کارایی عملی آن‌ها را در محیط‌های صنعتی کاهش می‌دهد[۱۱-۱۲].

با ظهور یادگیری عمیق و افزایش توان محاسباتی پردازنده‌های گرافیکی، تشخیص عیوب سطح فولاد وارد مرحله‌ای تازه شد. در حالی که روش‌های سنتی مبتنی بر ویژگی‌های دستی، در محیط‌های کنترل‌شده عملکرد قابل قبولی داشتند، اما در مواجهه با تغییرات نوری، نویز، و پیچیدگی بافت‌های سطحی عملکردشان افت محسوسی نشان می‌داد. این محدودیت‌ها زمینه‌ساز ورود شبکه‌های عصبی کانولوشنی به حوزه بازرسی بصری فولاد شد؛ شبکه‌هایی که با استخراج خودکار ویژگی‌ها از داده‌های خام، توانستند دقت، پایداری و تعمیم‌پذیری را بهبود دهند [۳].

در نخستین گام‌ها، پژوهشگران از معماری‌های کلاسیک نظیر AlexNet [۱۴]، VGG [۱۵] و ResNet [۱۶] برای طبقه‌بندی ساده عیوب استفاده کردند. در این روش ها به دنبال این بودند که شبکه‌ای طراحی کنند که با تأکید بر مدلهای از پیش آموزش دیده و استخراج ویژگی های سطح بالا، تنها با داده‌های محدود، انواع عیوب را با دقت و سرعت بالا شناسایی کنند. نتایج این پژوهش‌ها نشان داد که حتی مدل‌های سبک و کم‌پارامتر نیز قادرند در محیط‌های صنعتی با سرعت بالا، عملکردی پایدار و قابل‌اعتماد ارائه دهند [۱۷]. در سال‌های اخیر، شماری از پژوهش‌ها با هدف ارتقای دقت و پایداری مدل‌ها، به‌کارگیری رویکرد یادگیری ترکیبی[4] را مورد توجه قرار داده‌اند؛ رویکردی که در آن با تلفیق چند مدل یادگیری عمیق، تلاش می‌شود از قابلیت‌ها و ویژگی‌های مکمل هر مدل به‌صورت هم‌زمان بهره‌برداری شود [۱۸].

در ادامه، روند تحقیقات به‌سمت طراحی چارچوب‌های ترکیبی سوق یافت که بتوانند به‌طور هم‌زمان، طبقه‌بندی و مکان‌یابی دقیق عیوب سطحی را انجام دهند. در این رویکردها، با ترکیب شبکه‌های عصبی کانولوشنی و مدل‌های آشکارساز شیء مانند YOLO [۱۹]، ابتدا نوع عیب شناسایی شده و سپس موقعیت دقیق آن بر اساس نقشه‌های ویژگی استخراج‌شده تعیین می‌گردد. این روش گامی مؤثر در جهت توسعه سامانه‌های چندمرحله‌ای بازرسی بصری محسوب می‌شود که نه‌تنها قادر به تشخیص وجود عیب هستند، بلکه می‌توانند شکل، اندازه و موقعیت دقیق نقص را نیز با دقت در سطح پیکسل مشخص کنند[۲۰-۲۱].

در کنار این پیشرفت‌ها، برخی محققان متوجه شدند که تکیه صرف بر داده‌های برچسب‌خورده در محیط‌های واقعی امکان‌پذیر نیست؛ زیرا تولید چنین داده‌هایی در خطوط تولید صنعتی هزینه‌بر و زمان‌گیر است. این چالش موجب رشد روش‌های بدون‌نظارت شد. در این رویکرد، مدل‌ها با استفاده از تصاویر بدون عیب آموزش

[1] Shape Descriptors
[2] Contour-based descriptors
[3] Region-based descriptors
[4] Ensemble Learning

می‌بینند تا بتوانند هرگونه انحراف از الگوی نرمال را به‌عنوان عیب تشخیص دهند. از شبکه خودرمزگذار برای بازسازی تصویر سالم و مقایسه آن با تصویر واقعی استفاده کرد و اختلاف پیکسلی را نشانه وجود عیب و ناهنجاری در داده دانست [۲۲].

با این حال، هیچ‌یک از دو رویکرد نظارتی و بدون‌نظارت به‌تنهایی توان پاسخ‌گویی کامل به نیازهای پیچیده و پویای محیط‌های صنعتی را نداشتند. از این‌رو، نسل جدیدی از روش‌ها تحت عنوان یادگیری نیمه‌نظارتی[1] و یادگیری ضعیف‌نظارتی[2] مطرح شد. در این روش‌ها، حجم محدودی از داده‌های برچسب‌خورده با مجموعه‌ای بزرگ از داده‌های بدون برچسب ترکیب می‌شود تا مدل بتواند با بهره‌گیری از یادگیری ترکیبی، به دقتی نزدیک به روش‌های کاملاً نظارتی، اما با هزینه برچسب‌گذاری بسیار کمتر دست یابد. برای نمونه، یو و همکاران [۲۳] مدلی نیمه‌نظارتی مبتنی بر خودرمزگذارها[3] و شبکه‌های مولد[4] ارائه کردند که قادر بود ویژگی‌های ظریف و غیرخطی عیوب سطحی را به‌صورت خودکار استخراج کند. همچنین، ژانگ و همکاران در مدل CADN[۱۵]، تنها با استفاده از برچسب‌های سطح تصویر توانستند به‌طور هم‌زمان طبقه‌بندی و مکان‌یابی دقیق نقص را انجام دهند؛ رویکردی که چشم‌اندازی نوین برای توسعه سامانه‌های هوشمند بازرسی صنعتی فراهم ساخت و مسیر پژوهش‌های آینده را به‌سوی یادگیری داده‌محور و کم‌نمونه هموار کرد.

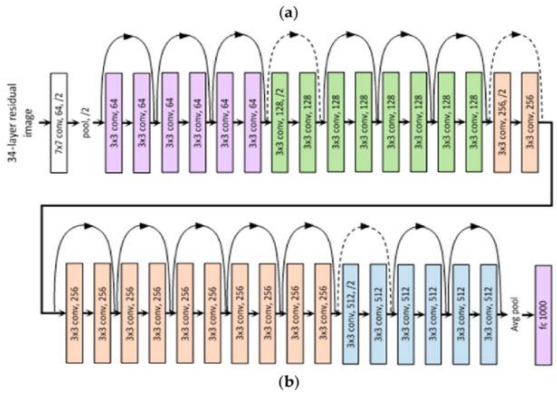
اغلب روش‌های موجود بروی داده‌ها و دسته‌بندی‌های شناخته‌شده از عیوب کارکرده و نوع خاصی مثل دوپوستگی تقریباً در هیچ کدام از مقالات موجود مورد بحث و بررسی قرار نگرفته است.

۳- مفاهیم پایه

در این مقاله به منظور تشخیص دو پوسته بودن از چندین شبکه شناخته شده استفاده‌شده است که در ادامه مختصری از آن‌ها بیان می‌شود.

۳-۱ ResNet

معماری ResNet با معرفی اتصالات باقیمانده[5] در شبکه‌های عصبی عمیق، تحولی بنیادین در حوزه یادگیری عمیق ایجاد کرد. این اتصالات به شبکه اجازه می‌دهند تا اطلاعات و گرادیان‌ها را از لایه‌های ابتدایی مستقیماً به لایه‌های عمیق‌تر منتقل کند و بدین ترتیب، مشکل ناپدید شدن گرادیان[6] را که یکی از چالش‌های اصلی در آموزش شبکه‌های بسیار عمیق است، برطرف سازد. در نتیجه، معماری ResNetامکان طراحی و آموزش شبکه‌هایی با عمق بسیار بالا را بدون افت کارایی فراهم می‌کند و به صورت چشمگیری عملکرد مدل در وظایف طبقه‌بندی تصویر را بهبود می‌بخشد [۱۶]. معماری این مدل در شکل (۱ نشان داده‌شده است.



شکل (۱): معماری مدل ResNet [۱۶]

۳-۲ EfficientNet

مدل EfficientNetB0 به‌عنوان نسخه پایه از خانواده EfficientNetدر سال ۲۰۱۹ توسط تیم گوگل معرفی شد. این معماری بر پایه مفهوم مقیاس‌گذاری ترکیبی طراحی‌شده است که در آن سه عامل اصلی شبکه عمق، عرض و رزولوشن ورودی به‌صورت هم‌زمان و متعادل افزایش می‌یابند. این رویکرد سبب می‌شود مدل با حداقل تعداد پارامتر، به حداکثر دقت ممکن در طبقه‌بندی تصاویر دست یابد. در نسخهB0 ، شبکه از بلوک‌های MBConv[7]استفاده می‌کند که پیش‌تر در معماری MobileNetV2معرفی‌شده بودند و به‌واسطه ساختار سبک و

[1] Semi-Supervised Learning
[2] Weakly-Supervised Learning
[3] Autoencoders
[4] Generative Networks

[5] Residual Connections
[6] Vanishing Gradient Problem
[7] Mobile Inverted Bottleneck Convolution

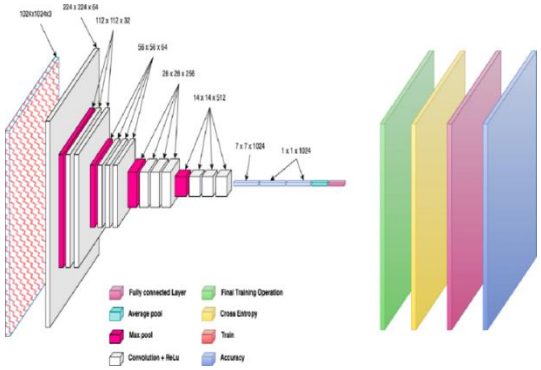
کارآمد خود، باعث کاهش قابل‌توجه حجم محاسبات می‌شوند افزون بر این، این مدل با بهره‌گیری از بهینه‌سازی خودکار معماری[1] و تنظیم دقیق هایپرپارامترها، تعادلی بهینه میان دقت، سرعت و مصرف منابع محاسباتی برقرار می‌سازد [۲۵] .

۳-۳ MobileNet

معماری MobileNet یکی از شبکه‌های عصبی پیچشی کارآمد و سبک‌وزن است که به‌طور خاص برای کاربردهای قابل‌حمل، سامانه‌های نهفته[2] و دستگاه‌های دارای منابع محاسباتی محدود طراحی‌شده است. نوآوری اصلی این معماری در به‌کارگیری پیچش‌های جداپذیر عمقی[3] نهفته است؛ رویکردی که عملیات پیچش متداول را به دو مرحله مجزا تقسیم می‌کند، به گونه‌ای که در مرحله نخست پیچش عمقی[4] هر فیلتر را به‌صورت مستقل بر روی یک کانال ورودی اعمال می‌کند و در مرحله دوم پیچش نقطه‌ای[5] با استفاده از فیلترهای ۱×۱ خروجی کانال‌ها را با یکدیگر ادغام می‌نماید. این طراحی موجب کاهش چشمگیر تعداد پارامترها و حجم محاسبات در مقایسه با شبکه‌های پیچشی کلاسیک می‌شود، در حالی‌که دقت مدل تا حد زیادی حفظ و حتی در برخی وظایف بهینه‌سازی می‌شود. بدین ترتیب، MobileNet یکی از گزینه‌های برتر برای استقرار مدل‌های یادگیری عمیق در محیط‌های عملیاتی با محدودیت منابع محاسباتی و توان پردازشی محسوب می‌گردد [۲۶]. معماری این مدل در شکل (۲ نشان داده‌شده است. در این پژوهش از MobileNetV2 که یک نسخه بهبودیافته از MobileNet است استفاده می‌کنیم.

۴-۳ لایه خروجی و تابع فعال‌سازی

برای مسئله طبقه‌بندی دودویی که در این پژوهش شامل تمایز بین نمونه‌های دوپوسته و غیردوپوسته است، از یک لایه تمام‌متصل در انتهای شبکه استفاده‌شده است. این لایه شامل یک نورون خروجی است که از تابع فعال‌سازی سیگموئید بهره می‌گیرد. خروجی این تابع مقداری بین ۰ و ۱ تولید می‌کند که می‌توان آن را به عنوان احتمال تعلق تصویر ورودی به کلاس دوپوسته بودن تفسیر کرد.



شکل (۲): معماری مدل MobileNet [۲۶]

۴- مجموعه داده‌ها

در این بخش، فرایندهای مرتبط با گردآوری و پیش‌پردازش داده‌ها که مبنای اصلی آموزش مدل‌های یادگیری عمیق در این پژوهش هستند، تشریح می‌گردند. از آنجا که عملکرد هر سامانه مبتنی بر یادگیری ماشین به‌طور مستقیم به کیفیت و کفایت داده‌های آموزشی وابسته است، اطمینان از صحت، تنوع و قابلیت اعتماد داده‌ها از اهمیت ویژه‌ای برخوردار است. بدین منظور، مراحل گردآوری داده‌ها با هدف دستیابی به تصاویری دقیق، باکیفیت و منطبق بر شرایط واقعی خط تولید طراحی‌شده‌اند. سپس مجموعه‌ای از فرایندهای پیش‌پردازش بر روی داده‌ها اعمال گردیده است تا ورودی‌های شبکه‌های عصبی از ثبات و یکنواختی لازم برخوردار باشند. این اقدامات، زمینه‌ساز افزایش کارایی مدل در تشخیص عیوب سطحی فولاد و بهبود قابلیت تعمیم آن در محیط‌های صنعتی واقعی است.

۴-۱ فرایند گردآوری داده‌ها

به‌منظور ایجاد مجموعه‌داده‌ای قابل‌اعتماد برای آموزش مدل‌های یادگیری عمیق، تصاویر میلگردهای فولادی در حین خروج از یک محفظه صنعتی (خط‌نورد) توسط دوربین‌های صنعتی Basler با دقت بالا ثبت گردیدند (شکل (۳). انتخاب این دوربین‌ها به‌دلیل رزولوشن بالا، نرخ فریم مناسب و پایداری عملکرد در شرایط سخت محیطی صورت گرفته است، ویژگی‌هایی که آن‌ها را به

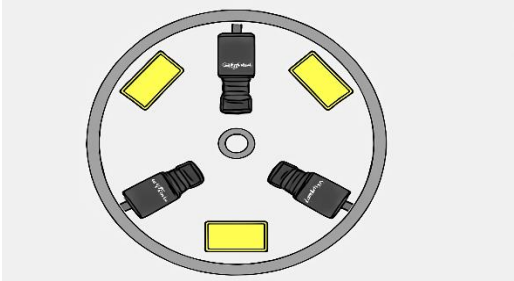
[1] Neural Architecture Search
[2] Embedded Systems
[3] Depthwise Separable Convolutions
[4] Depthwise Convolution
[5] Pointwise Convolution

گزینه‌ای ایده‌آل برای کاربردهای بینایی ماشین در محیط‌های صنعتی تبدیل می‌کند.

شکل (۳): دوربین Basler مورد استفاده برای جمع‌آوری تصاویر

به منظور دستیابی به پوشش کامل سطح میلگرد و ثبت تصاویر با بیشترین جزئیات، از سه دوربین صنعتی Basler بر روی یک محفظه استوانه‌ایی استفاده‌شده است. این دوربین‌ها در زاویه‌های ۱۲۰ درجه نسبت به یکدیگر نصب گردیده‌اند تا نمایی جامع از کل محیط پیرامونی میلگرد فراهم شود و هیچ نقطه‌ای از سطح آن خارج از میدان دید قرار نگیرد. در مرحله تشخیص، اگر مدل یادگیری عمیق در هر یک از سه تصویر ثبت‌شده وجود عیب دوپوستگی را شناسایی کند، نمونه مربوطه به عنوان میلگرد معیوب طبقه‌بندی می‌شود.

برای اطمینان از ثبات شرایط نوری و یکنواختی روشنایی در فرآیند تصویربرداری، از چهار لامپ زنون صنعتی در اطراف محفظه استفاده‌شده است تا یک محیط ایستا و با نور کنترل‌شده ایجاد شود. این تنظیمات موجب کاهش نویز نوری و بهبود کیفیت ویژگی‌های استخراج‌شده توسط مدل می‌گردد .آرایش کلی اجزای سیستم تصویربرداری به صورت شماتیک در شکل (٤ نشان داده‌شده است.

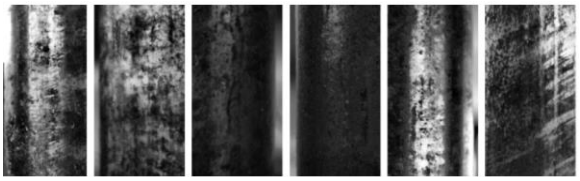


شکل (٤): نحوه چینش دوربین و چراغ‌ها در محیط

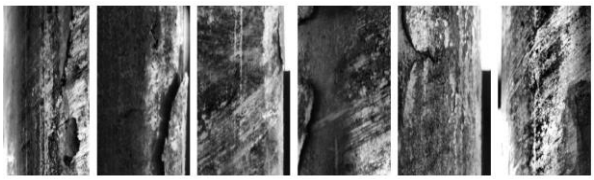
بر همین اساس، ویژگی‌های مجموعه تصاویر گردآوری‌شده به‌صورت زیر خلاصه می‌شود:

- **تصاویر سیاه و سفید:** تمامی تصاویر به صورت تک کاناله ثبت‌شده‌اند. این رویکرد موجب کاهش حجم داده‌ها و بار محاسباتی در مراحل بعدی پردازش می‌شود.
- **تصویربرداری در حین حرکت**: از آن‌جا که شمش یا میلگرد در حین فرآیند تولید در حال حرکت است، تصاویر متوالی در طول مسیر حرکت ثبت می‌شوند تا پوشش کامل سطح میلگرد تضمین‌شده و احتمال عدم ثبت نواحی دارای عیب به حداقل برسد.
- **برچسب‌گذاری توسط متخصصان**: پس از مرحله تصویربرداری، تصاویر توسط متخصصان مجرب در حوزه متالورژی و بازرسی فولاد مورد بازبینی دقیق قرارگرفته و به صورت دستی برچسب‌گذاری می‌شوند. در این فرآیند، هر تصویر در یکی از دو کلاس دوپوسته (معیوب) یا غیردوپوسته (سالم) قرار می‌گیرد. این مرحله برای ایجاد مجموعه‌داده‌ای آموزشی و آزمایشی با کیفیت بالا، که شرط لازم برای آموزش مؤثر مدل‌های یادگیری عمیق است، اهمیت حیاتی دارد.

نمونه‌هایی از تصاویر سالم و معیوب در شکل‌های ۵ و ٦ نمایش داده‌شده‌اند.



شکل (۵): نمونه تصویر میلگرد فولادی سالم.



شکل (٦): نمونه تصویر میلگرد فولادی با عیب دو پوسته.

این مجموعه داده به دو بخش آموزش و آزمون تقسیم‌شده است. جزئیات آماری آن در جدول (۱ آورده شده است. در این تقسیم‌بندی، کلاس صفر برای تصاویر سالم و کلاس یک برای تصاویر معیوب در نظر گرفته‌شده است.

جدول (۱): توزیع داده‌های آموزشی و آزمون

بخش داده	کلاس سالم	کلاس دوپوسته (معیوب)	جمع کل
آموزش	۱۹۳٤	۱٦۳٦	۳۵۷۰
آزمون	۲۰۰	۲۰۰	٤۰۰

همان‌گونه که مشاهده می‌شود، مجموعه‌داده تا حدودی نامتوازن است (تعداد تصاویر سالم اندکی بیشتر است)، اما این عدم توازن در حدی نیست که نیاز به اعمال روش‌های توازن‌دهی یا وزن‌دهی خاص در تابع زیان داشته باشد.

٤-۲ پیش‌پردازش تصاویر

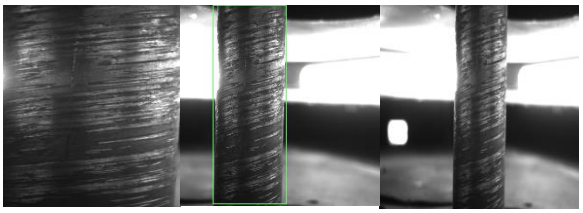
پیش‌پردازش تصاویر مرحله‌ای کلیدی در آماده‌سازی داده‌ها پیش از آموزش مدل است، که هدف آن بهبود کیفیت تصاویر، حذف نویز و استانداردسازی ورودی‌ها برای شبکه‌های عصبی است. در این پژوهش مراحل زیر اجرا شده‌اند:

٤-۲-۱ تشخیص لبه یا قطعه‌بندی میلگرد[1]

در ابتدا، بخش مربوط به میلگرد باید از پس‌زمینه جدا شود. این کار می‌تواند از طریق روش‌های تشخیص لبه (مانند فیلتر سوبل، کنی یا لاپلاس) [٤] یا تکنیک‌های قطعه‌بندی تصویر (مانند آستانه‌گذاری، قطعه‌بندی مبتنی بر رشد ناحیه یا شبکه‌های عصبی قطعه‌بندی مانند U-Net) انجام شود [۵]. در اینجا هدف این است که تنها ناحیه مورد نظر که شامل سطح فولاد است، حفظ شود و اطلاعات غیرضروری پس‌زمینه حذف گردد. این کار باعث تمرکز مدل بر ویژگی‌های مرتبط با عیب می‌شود. در شکل (۷ نمونه‌ای از تصویر ورودی و تصویر قطعه‌بندی شده آورده شده است.

برای جداسازی میلگرد از پس‌زمینه، ابتدا یک تصویر مرجع[2] از محیط بدون حضور میلگرد توسط دوربین‌ها ثبت و در حافظه ذخیره می‌شود. در ادامه، هر تصویر جدیدی که توسط دوربین‌ها دریافت می‌گردد با تصویر مرجع مقایسه شده و اختلاف پیکسل‌ها محاسبه می‌شود. این فرایند سبب حذف اجزای ثابت محیط و برجسته‌سازی نواحی متغیر، یعنی همان میلگرد، خواهد شد. پس از این مرحله، برای استخراج دقیق مرزهای میلگرد از الگوریتم تشخیص لبه سوبل استفاده می‌شود. این الگوریتم تغییرات شدت روشنایی را در تصویر محاسبه کرده و خطوط و لبه‌های میلگرد را به‌وضوح نمایان می‌سازد.

در صورتی که در خروجی الگوریتم سوبل هیچ لبه‌ای شناسایی نشود، نتیجه گرفته می‌شود که میلگرد هنوز وارد ناحیه دید دوربین نشده است. در این حالت، تصویر مربوطه به عنوان تصویر محیط شناخته شده و برای مدل پردازشی ارسال نمی‌شود. به این ترتیب از پردازش غیرضروری جلوگیری شده و منابع محاسباتی سیستم به‌صورت بهینه مصرف می‌شوند. اما در صورتی که لبه‌ها شناسایی شوند، یک ماسک دودویی از نواحی شناسایی‌شده تولید می‌گردد. این ماسک بر روی تصویر اصلی اعمال‌شده و تنها نواحی مربوط به میلگرد باقی می‌مانند. سپس با استفاده از این ماسک، تصویر برش خورده به عنوان نسخه نهایی میلگرد بدون پس‌زمینه استخراج می‌شود. این تصویر برش‌خورده به عنوان ورودی به مراحل بعدی پردازش ارسال خواهد شد.



الف ب ج

شکل (۷): الف) تصویر ورودی ب) ماسک تشخیص میلگرد ج) تصویر قطعه‌بندی شده

٤-۲-۲ تغییر اندازه تصویر[3]

پس از جداسازی میلگرد، تصاویر به اندازه یکنواخت ۲۲٤×۲۲٤ پیکسل تغییر اندازه داده می‌شوند. این اندازه استاندارد ورودی برای بسیاری از مدل‌های پیش‌آموزش‌دیده شبکه‌های عصبی پیچشی (مانند VGG, ResNet, Inception) است [٦]. تغییر اندازه یکنواخت، تضمین می‌کند که تمامی تصاویر ورودی به مدل دارای ابعاد یکسان باشند.

٤-۲-۳ نرمال‌سازی[4]

نرمال‌سازی یکی از مهم‌ترین مراحل پیش‌پردازش برای شبکه‌های عصبی است. این فرآیند مقادیر شدت پیکسل‌ها را به یک محدوده

[1] Segmentation
[2] Background Image
[3] Resizing
[4] Normalization

استاندارد مقیاس‌بندی می‌کند [۷]. برای تصاویر سیاه و سفید، معمولاً مقادیر پیکسل‌ها در محدوده [۰،۲۵۵] هستند. نرمال‌سازی با تقسیم هر مقدار پیکسل بر ۲۵۵، آن‌ها را به محدوده [۰،۱] می‌آورد. این کار به همگرایی سریع‌تر و پایدارتر مدل در طول آموزش کمک می‌کند و از تأثیر مقیاس‌های مختلف ویژگی‌ها بر فرآیند یادگیری جلوگیری می‌کند.

۴-۲-۴ داده‌افزایی[1]

برای افزایش حجم و تنوع مجموعه داده آموزشی و جلوگیری از بیش‌برازش[2]، از تکنیک داده‌افزایی استفاده می‌شود [۱۶]. در این روش، با اعمال تبدیلات هندسی و نوری بر روی تصاویر، نمونه‌های جدیدی تولید می‌شود بدون آنکه ماهیت یا برچسب آن‌ها تغییر کند. تبدیلات مورد استفاده در این پژوهش عبارت‌اند از:

- **چرخش**[3]: چرخش تصاویر در زوایای مختلف.
- **قرینه‌سازی**[4]: قرینه‌سازی افقی یا عمودی تصاویر.
- **بزرگ‌نمایی**[5]: بزرگ‌نمایی یا کوچک‌نمایی بخش‌هایی از تصویر.
- **تغییر روشنایی و کنتراست**[6]: ایجاد تغییرات جزئی در روشنایی و کنتراست.
- **برش تصادفی**[7]: حذف تصادفی بخش‌های غیرضروری تصویر.
- **افزودن نویز**[8]: افزودن نویز گوسی یا نمکی به تصویر.

اجرای این تبدیلات موجب می‌شود مدل نسبت به تغییرات جزئی در داده‌های ورودی مقاوم‌تر شده و قابلیت تعمیم‌پذیری آن افزایش یابد.

۵- روش پیشنهادی

پس از انجام مراحل پیش‌پردازش، تصاویر آماده ورود به مدل‌های شبکه‌های عصبی مصنوعی برای انجام فرآیند طبقه‌بندی هستند. در این پژوهش، از شبکه‌های عصبی پیچشی[9] استفاده‌شده است که به‌طور ویژه برای پردازش داده‌های تصویری طراحی گردیده‌اند و در وظایف مرتبط با طبقه‌بندی تصویر عملکردی بسیار کارآمد و دقیق از خود نشان داده‌اند.

معماری‌های گوناگونی از شبکه‌های CNN وجود دارند که می‌توانند برای شناسایی عیوب سطحی مورد استفاده قرار گیرند. در این پژوهش، با هدف تشخیص خودکار عیب دوپوستگی در میلگردهای فولادی، مجموعه‌ای از معماری‌های شبکه‌های عصبی پیچشی مورد استفاده و ارزیابی قرارگرفته‌اند. معماری‌های منتخب شامل سه شبکه پیش‌آموزش‌دیده یعنی MobileNetV2، ResNet50 و EfficientNetB0 و یک شبکه Custom CNN طراحی‌شده از پایه می‌باشند. این ترکیب، امکان مقایسه دقیق میان مدل‌های سبک و عمیق، و نیز ارزیابی تأثیر پیش‌آموزش در مقابل آموزش از صفر را فراهم می‌سازد.

۵-۱ مدل پیش‌آموزش‌دیده

مدل‌هایMobileNeV2 ، ResNet50و EfficientNetB0 که پیش‌تر بر روی مجموعه داده بزرگ ImageNet آموزش‌دیده‌اند، در این پژوهش به‌عنوان استخراج کننده‌های ویژگی مورد استفاده قرار گرفتند. به‌منظور سازگاری این مدل‌ها با مسئله حاضر، دو لایه انتهایی هر شبکه آزادشده[10] تا در فرآیند تنظیم دقیق[11]، آموزش یابند، درحالی‌که سایر لایه‌ها ثابت[12] باقی مانده‌اند تا ویژگی‌های عمومی حاصل از مرحله پیش‌آموزش حفظ گردند.

پس از بخش استخراج ویژگی، یک لایه تمام‌متصل با ٦٤ نورون و تابع فعال‌سازی ReLUاضافه شد تا ویژگی‌های استخراج‌شده را ترکیب کند. در انتها، یک لایه خروجی با یک نورون و تابع فعال‌سازی سیگموئید برای انجام طبقه‌بندی دودویی افزوده گردید. این طراحی امکان بهره‌گیری هم‌زمان از دانش ازپیش‌آموخته شبکه و سازگاری آن با داده‌های صنعتی محدود را فراهم می‌کند و به بهبود دقت، پایداری و سرعت همگرایی مدل منجر می‌شود.

[1]Data Augmentation
[2]Overfitting
[3]Rotation
[4]Flipping
[5]Zooming
[6]Brightness and Contrast Adjustment
[7]Random Cropping
[8]Adding Noise
[9]Convolutional Neural Networks(CNNs)
[10]Unfrozen
[11]Fine-Tuning
[12]Frozen

۵-۲ مدل پیچشی ساده

در کنار مدل‌های پیش‌آموزش‌دیده، یک شبکه Custom CNNنیز به‌صورت مستقل و بدون استفاده از وزن‌های اولیه طراحی و از ابتدا آموزش داده شد تا عملکرد مدل‌های مبتنی بر یادگیری از صفر نیز ارزیابی گردد.

این معماری شامل سه بلوک پیچشی متوالی است که هر بلوک از یک لایه کانولوشن دوبعدی و یک لایه ادغام حداکثری تشکیل‌شده است. تعداد فیلترها در لایه‌های پیچشی به‌ترتیب ۳۲، ٦٤ و ۱۲۸ در نظر گرفته‌شده است تا با افزایش عمق شبکه، قابلیت استخراج ویژگی‌های پیچیده‌تر از داده‌های تصویری فراهم شود.

خروجی لایه‌های پیچشی پس از تبدیل به بردار مسطح به یک لایه تمام‌متصل با ۱۲۸ نورون و تابع فعال‌سازی ReLU متصل می‌شود. به‌منظور کاهش خطر بیش‌برازش، یک لایه Dropout با نرخ ۰,٦ به‌کار گرفته شد. در نهایت، مشابه سایر مدل‌ها، یک لایه خروجی سیگموئید برای طبقه‌بندی دودویی اضافه گردید. در تحلیل‌ها، این مدل با نماد Custom_CNNنمایش داده‌شده است.

۵-۳ آموزش و ارزیابی مدل

آموزش مدل‌های یادگیری عمیق مستلزم استفاده از مجموعه‌داده‌ای بزرگ، متنوع و با کیفیت بالا است. علاوه بر این، ارزیابی دقیق عملکرد مدل برای اطمینان از قابلیت تعمیم آن به داده‌های جدید از اهمیت ویژه‌ای برخوردار است. در این بخش، جزئیات مربوط به مجموعه‌داده، داده‌افزایی و تنظیمات آموزش، معیارهای ارزیابی مدل‌ها تشریح می‌شود.

۵-۳-۱ تنظیمات آموزش

فرآیند آموزش مدل‌های یادگیری عمیق مستلزم تعیین مجموعه‌ای از هایپرپارامترها است که به‌صورت مستقیم بر سرعت همگرایی و دقت نهایی مدل تأثیر می‌گذارند. در این پژوهش، مقادیر انتخاب‌شده برای هایپرپارامترهای اصلی آموزش در جدول (۲) ارائه‌شده است.

جدول (۲): تنظیمات آموزش مدل‌های یادگیری عمیق.

پارامتر	مقدار
ابعاد ورودی تصویر	۲۲٤×۲۲٤ پیکسل
اندازه دسته	۳۲
تعداد دوره‌های آموزشی	۵۰
نرخ یادگیری	۰/۰۰۰۰۰۵
نرخDropout	۰/٦
توقف زودهنگام	پس از ۱۰ گام بدون بهبود
بهینه‌ساز	Adam
تابع زیان	آنتروپی متقاطع دودویی[1]

تابع زیان آنتروپی متقاطع دودویی برای طبقه‌بندی دودویی به‌کار گرفته‌شده است، زیرا اختلاف بین احتمال پیش‌بینی‌شده توسط مدل و برچسب واقعی را به‌صورت مستقیم اندازه‌گیری می‌کند. بهینه‌ساز Adamنیز به دلیل سرعت همگرایی بالا و پایداری در آموزش شبکه‌های عمیق انتخاب‌شده است. همچنین از زمان‌بندی نرخ یادگیری تطبیقی استفاده‌شده تا در مراحل پایانی آموزش، نرخ یادگیری به‌تدریج کاهش یابد و مدل به حداقل بهینه همگرا شود.

به منظور پیاده‌سازی سیستم جاری از زبان برنامه‌نویسی پایتون و چارچوب Tensorflow استفاده‌شده است. همچنین برای توسعه و اجرای کدها از Googel Colab با ۱۵ گیگابایت GPU و ۱٦ گیگابایت رم استفاده‌شده است. از لحاظ زمانی، اجرای مدل‌ها از یک ساعت تا ۲ ساعت متفاوت بوده است و زمان اجرای کل مدل‌ها، حدود ۸ ساعت به طول انجامید.

۵-۳-۲ معیارهای ارزیابی

برای ارزیابی عملکرد مدل‌های طبقه‌بندی و سنجش دقت آن‌ها در تشخیص میلگردهای معیوب، از مجموعه‌ای از معیارهای ارزیابی متداول در حوزه یادگیری ماشین استفاده‌شده است. این معیارها شامل موارد زیر هستند:

- **دقت**[2]: نسبت کل پیش‌بینی‌های صحیح به تعداد کل نمونه‌ها. در معادله ۱ فرمول آن نشان داده‌شده است.

[1] Binary Cross-Entropy

[2] Accuracy

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (۱)$$

- **صحت**[1]: نسبت مثبت‌های واقعی به کل پیش‌بینی‌های مثبت. در معادله ۲ فرمول آن نشان داده‌شده است.

$$Precision = \frac{TP}{TP+FP} \quad (۲)$$

- **یادآوری**[2]: نسبت نمونه‌های مثبت واقعی که به‌درستی توسط مدل شناسایی شده‌اند به کل نمونه‌های مثبت واقعی موجود در داده. در معادله ۳ فرمول آن نشان داده‌شده است.

$$Recall = \frac{TP}{TP+FN} \quad (۳)$$

- **امتیاز F1 (F1-Score)**: میانگین هارمونیک دقت و فراخوانی، که به ویژه برای مجموعه داده‌های نامتوازن معیار مناسبی به‌شمار می‌رود. در معادله ٤ فرمول آن نشان داده‌شده است.

$$F1-Score = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (٤)$$

- **منحنی ROC[3] و مساحت زیر منحنی**[4]: منحنی ROC رابطه بین نرخ مثبت واقعی و نرخ مثبت کاذب را در آستانه‌های مختلف طبقه‌بندی نشان می‌دهد و مقدار AUC نشان‌دهنده توانایی کلی مدل در تفکیک کلاس‌های مثبت و منفی است و هرچه این مقدار به ۱ نزدیک‌تر باشد، عملکرد مدل مطلوب‌تر است [۲۱].

٦- نمودارها و نتایج

در این بخش، نتایج حاصل از آموزش مدل‌های مختلف شبکه عصبی پیچشی و ارزیابی عملکرد آن‌ها بر اساس معیارهای استاندارد ازجمله دقت، صحت، یادآوری، امتیاز F1 و مساحت زیر منحنی (AUC) ارائه می‌شود. علاوه بر این، نمودارهای مربوط به فرآیند آموزش و ارزیابی مدل‌ها برای تحلیل بهتر رفتار آن‌ها در طول آموزش بررسی شده‌اند.

٦-١ نتایج کمی

در جدول ۳ مقادیر نهایی معیارهای ارزیابی برای مدل‌های مختلف آورده شده است. مطابق نتایج مشاهده‌شده، مدل MobileNetV2 بالاترین عملکرد را در میان مدل‌های بررسی‌شده داشته و به دقت کلی ۹۱/۲۵ و مقدار AUC برابر با ۰/۹۷۳۱ دست یافته است. مدل Custom_CNN نیز با اختلاف اندک در جایگاه دوم قرارگرفته است، در حالی‌که عملکرد مدل‌های ResNet50 و EfficientNetB0 پایین‌تر بوده است.

جدول (۳): نتایج ارزیابی مدل‌ها بر اساس معیارهای مختلف

مدل	دقت	صحت	یادآوری	امتیاز F1	AUC
ResNet50	۰/۷۳۷۵	۰/۷۳۸۰	۰/۷۳۷۵	۰/۷۳۷٤	۰/۸۱۲۰
EfficientNetB0	۰/۵	۰/۲۵۰۰	۰/۵۰۰۰	۰/۳۳۳۳	۰/۲۹۱٦
MobileNetV2	۰/۹۱۲۵	۰/۹۱٦۳	۰/۹۱۲۵	۰/۹۱۲۳	۰/۹۷۳۱
Custom_CNN	۰/۹	۰/۹۰۲٦	۰/۹۰۰۰	۰/۸۹۹۸	۰/۹٦۷۵

در مقایسه عملکرد مدل‌ها، نتایج نشان داد که دو مدل ResNet50 و EfficientNetB0دقت پایین‌تری نسبت به سایر مدل‌ها دارند. در مدل EfficientNetB0، عملکرد ضعیف (دقت ۰/۵۰ و مقدار AUCبرابر با ۰/۲۹) عمدتاً ناشی از حساسیت بالا به تنظیمات ورودی، اندازه تصویر و محدودیت حجم داده‌ها است. این معماری که بر مبنای مقیاس‌گذاری ترکیبی طراحی‌شده، برای دستیابی به تعمیم مناسب نیازمند داده‌های متنوع‌تر و تنظیم دقیق پارامترهای یادگیری است. به همین دلیل، مدل دچار کم‌برازش شده و در تفکیک صحیح نمونه‌های سالم و معیوب ناکام مانده است. در مقابل، مدل ResNet50با وجود دقت بالاتر (۰/۷۳۷۵) نسبت به EfficientNetB0، در مقایسه با مدل‌های سبک‌تر عملکرد ضعیف‌تری نشان داده است. دلیل اصلی این مسئله، پیچیدگی زیاد و تعداد بالای پارامترها در شرایطی است که حجم داده‌ها محدود

[1]Precision

[2]Recall

[3]Receiver Operating Characteristic

[4]Area Under Curve(AUC)

بوده و مدل مستعد بیش‌برازش شده است. علاوه بر این، ResNet50که برای تصاویر طبیعی آموزش‌دیده، در استخراج ویژگی‌های ظریف و بافت‌محور تصاویر صنعتی فولاد کارایی کمتری دارد. به‌طور کلی، می‌توان نتیجه گرفت که در داده‌های صنعتی با ویژگی‌های محلی و نویز بالا، مدل‌های سبک‌تر مانند MobileNetو شبکه پیشنهادی سفارشی Custom_CNN به دلیل انعطاف بیشتر و نیاز کمتر به داده‌های حجیم، توانایی بالاتری در یادگیری الگوهای واقعی و تعمیم به داده‌های جدید دارند.

۶-۲ پیچیدگی مدل‌ها

برای ارزیابی کارایی محاسباتی، تعداد پارامترهای قابل آموزش هر مدل در جدول ٤ ارائه‌شده است. با وجود آنکه مدل Custom_CNNبیشترین تعداد پارامترها را دارد (حدود ۱۱ میلیون)، مدل MobileNetبا تنها حدود ١٦٤ هزار پارامتر، عملکرد بهتری در تمامی معیارها داشته است. این موضوع نشان می‌دهد که MobileNetدر عین سادگی و کم‌حجمی، از لحاظ کارایی محاسباتی بسیار بهینه‌تر است.

جدول (٤): تعداد پارامترهای قابل آموزش مدل‌های مختلف

مدل	تعداد پارامترهای قابل آموزش
ResNet50	٤٠١,٢٦٢
EfficientNetB0	٣٤٦,٢٩٠
MobileNetV2	٠٩٧,١٦٤
Custom_CNN	٠٨٩,١٦٩,١١

۶-۳ تحلیل کیفی و بصری عملکرد مدل‌ها

برای درک بهتر رفتار مدل‌ها، ماتریس‌های درهم‌ریختگی[1] و منحنی‌های ROC برای تمامی مدل‌ها رسم شده‌اند. ماتریس‌های درهم‌ریختگی نشان می‌دهند که مدل MobileNet و Custom_CNNبیشترین نرخ پیش‌بینی صحیح را در هر دو کلاس سالم و معیوب دارند. در مقابل، مدل EfficientNetB0 عملکرد ضعیفی در تمایز بین کلاس‌ها از خود نشان داده است. عملکرد ضعیف مدل EfficientNetB0در ماتریس درهم‌ریختگی نشان می‌دهد که این شبکه در تفکیک صحیح نمونه‌های سالم و معیوب دچار خطاهای متعدد شده و نتوانسته ویژگی‌های بافتی مرتبط با عیب دوپوستگی را به‌درستی یاد بگیرد. این مسئله عمدتاً ناشی از کم‌برازش به‌دلیل حجم محدود داده‌ها، حساسیت بالا به تنظیمات ورودی و مقیاس تصویر، و ناتوانی در استخراج ویژگی‌های موضعی ظریف سطح فولاد است. علاوه بر این، وابستگی مدل به وزن‌های پیش‌آموزش‌یافته مبتنی بر داده‌های دنیای واقعی موجب شده که در تطبیق با داده‌های صنعتی خاکستری عملکرد ضعیف‌تری نشان دهد، به‌گونه‌ای که توزیع پیش‌بینی‌ها در ماتریس درهم‌ریختگی تقریباً تصادفی مشاهده می‌شود. در شکل (۸ ماتریس‌های درهم‌ریختگی مربوط به مدل‌های مختلف نمایش داده‌شده‌اند که میزان دقت مدل‌ها را در تشخیص صحیح تصاویر سالم و دارای عیب دوپوستگی نشان می‌دهند. (کلاس صفر سالم، کلاس یک ناسالم)

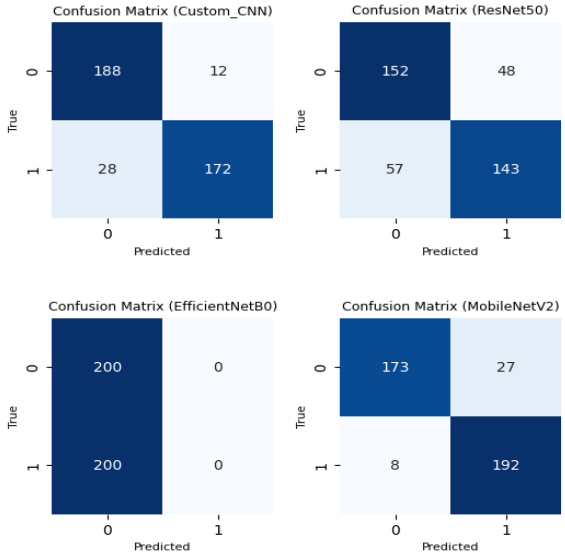
در **Error! Reference source not found.**، نتایج حاصل از نمودارهای ROCمدل‌های مورد بررسی نشان می‌دهد که میزان تفکیک‌پذیری میان دو کلاس سالم و معیوب در مدل‌های مختلف متفاوت بوده است. همان‌گونه که مقادیر AUCبیان می‌کنند، مدل‌های MobileNetو Custom_CNNبا مقادیر ۹۷/۳۱ و ۹٦/۷٥ درصد، بالاترین توانایی در تمایز صحیح بین نمونه‌های مثبت و منفی را داشته‌اند. این امر نشان‌دهنده پایداری بالا و حساسیت مناسب این دو مدل در تشخیص عیب دوپوستگی است. در مقابل، مدل ResNet50با مقدار ۸۱/۲ درصد عملکرد متوسطی داشته و مدل EfficientNetB0با مقدار بسیار پایین ۲۹/۱٦ درصد نشان داده است که توان تمایز بسیار ضعیفی دارد و عملاً خروجی آن نزدیک به حد تصادف است.

شکل منحنی ROC مدل EfficientNetB0تقریباً در امتداد قطر پایین نمودار قرار دارد که بیانگر ناتوانی این مدل در تفکیک مؤثر نمونه‌های سالم (کلاس ۰) از نمونه‌های معیوب (کلاس ۱) است. این رفتار در بیشتر موارد ناشی از دو عامل فنی کلیدی است. نخست، هم‌پوشانی زیاد در توزیع احتمالات خروجی دو کلاس که سبب شده مدل نتواند مرز تصمیم‌گیری روشنی میان نمونه‌های سالم و معیوب ایجاد کند؛ در نتیجه، مقادیر خروجی برای هر دو

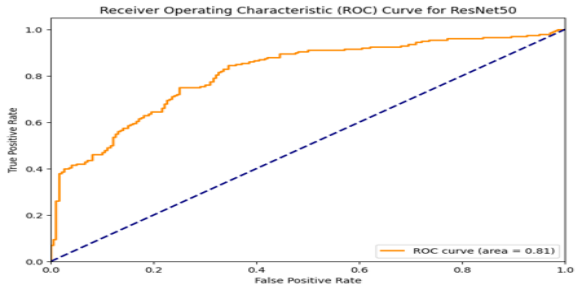
[1] Confusion Matrices

کلاس در ناحیه‌ای مشابه متمرکزشده و قدرت تفکیک‌پذیری مدل کاهش‌یافته است. دوم، اشباع تابع فعال‌سازی سیگموئید در لایه خروجی که منجر به تمرکز مقادیر پیش‌بینی در بازه میانی (۰/٤ تا ۰/٦) گردیده است؛ این امر بیانگر آن است که مدل در بیشتر موارد احتمال مشابهی برای هر دو کلاس تولید کرده و نسبت به تغییر آستانه‌های تصمیم‌گیری حساسیت کافی ندارد. چنین رفتاری باعث می‌شود نرخ مثبت واقعی و نرخ مثبت کاذب به‌صورت هم‌زمان افزایش یابد و در نتیجه، منحنی ROC به جای صعود، در امتداد قطر قرار گیرد که نشانگر افت محسوس در مقدار AUC و عملکرد ضعیف مدل در جداسازی کلاس‌ها است.

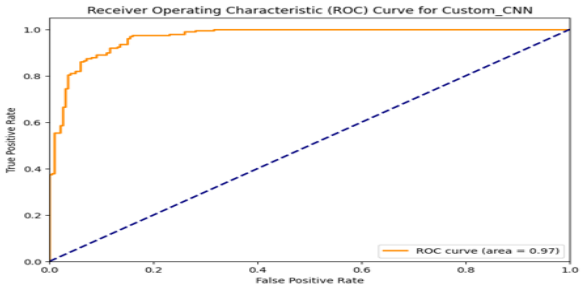
به‌طور کلی، تحلیل منحنی‌های ROC تأیید می‌کند که مدل‌های سبک‌تر و بهینه‌تر از نظر ساختار، در داده‌های صنعتی با نویز و تغییرات نوری بالا، نسبت به مدل‌های عمیق و پیچیده عملکرد دقیق‌تر و پایدارتری دارند.



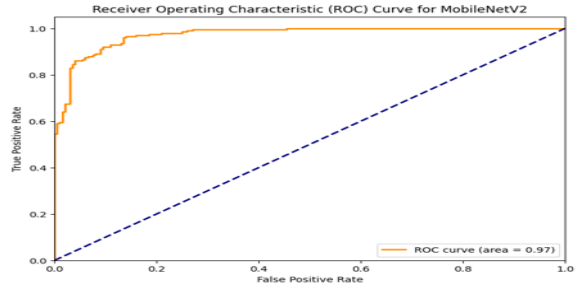
شکل (۸): ماتریس‌های درهم ریختگی مدل‌های مورد بررسی



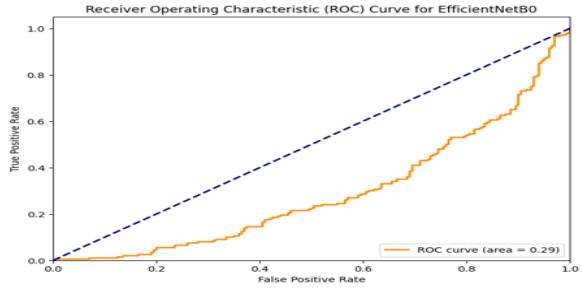
ب) منحنی ROC برای مدلResenet



الف) منحنی ROC برای مدل پیچشی ساده



د) منحنی ROC برای مدلMobileNet



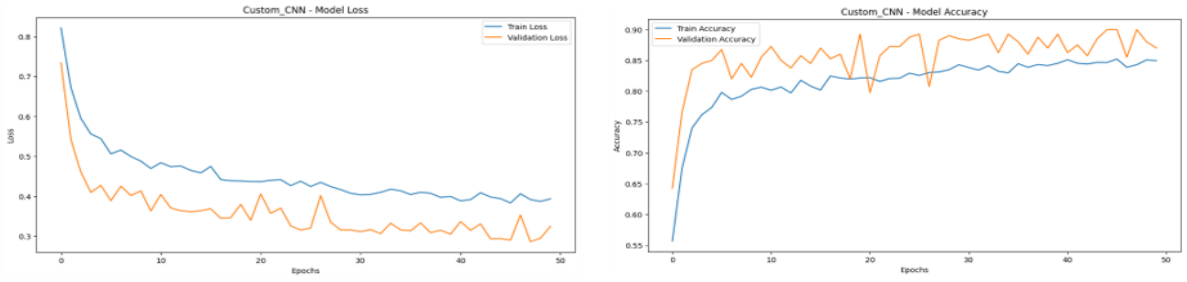
ج) منحنی ROC برای مدل EfficientNetB0

شکل (۹): منحنی‌های ROC مدل‌های مورد بررسی

٤-٦ تحلیل فرایند آموزش و همگرایی

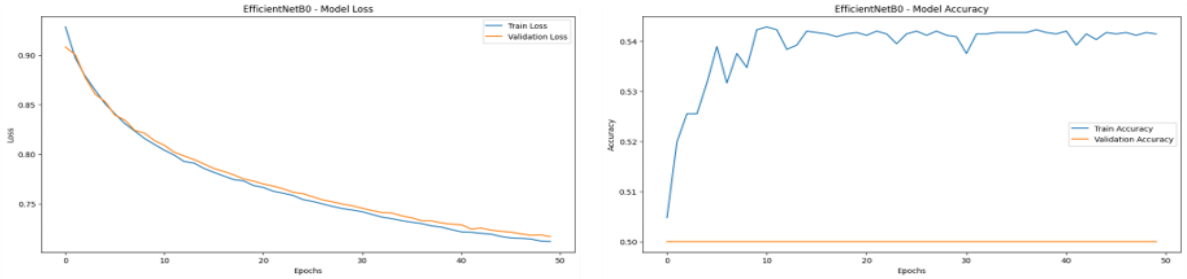
نمودارهای دقت و زیان مربوط به مراحل آموزش و اعتبارسنجی مدل‌ها در شکل‌های ۱۰ تا ۱۳ ارائه‌شده‌اند. بررسی این نمودارها نشان می‌دهد:

- مدل‌های ResNet50 و Custom_CNN دارای نوسانات آموزشی قابل‌توجهی هستند که ناشی از تعداد زیاد پارامترهای قابل یادگیری است.
- مدل MobileNet روند همگرایی یکنواخت و پایداری دارد که بیانگر تعادل مناسب میان دقت و سادگی معماری است.
- مدل EfficientNetB0 دچار کم‌برازش شده و از یادگیری مؤثر داده‌ها بازمانده است.



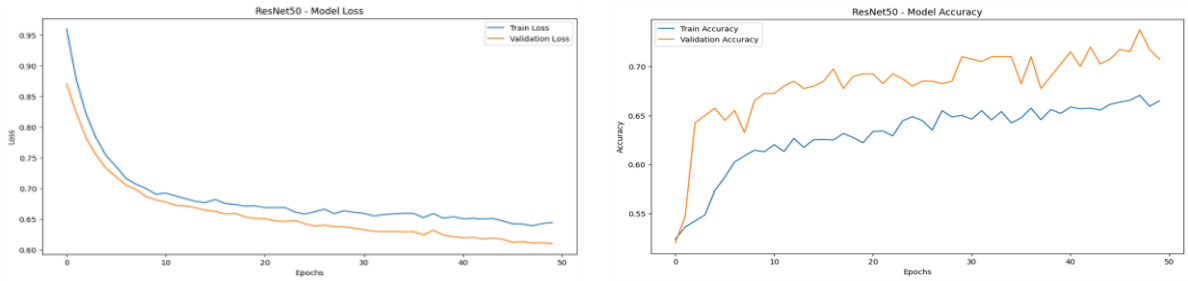
ب) خطا الف) دقت

شکل (۱۰): نمودار دقت و زیان آموزش و اعتبارسنجی مدل پیچشی ساده



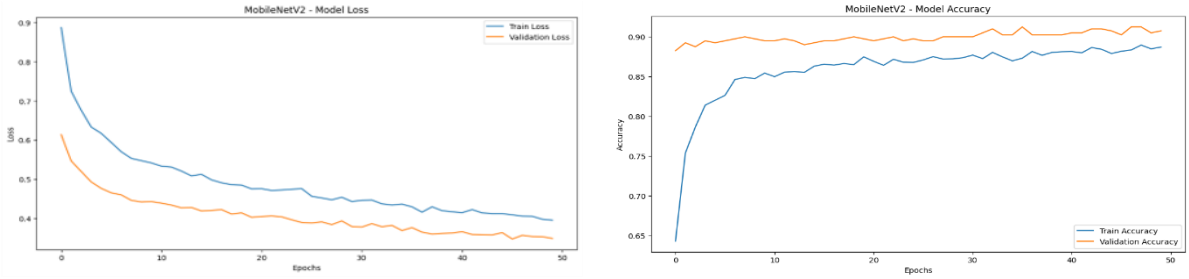
ب) خطا الف) دقت

شکل (۱۱): نمودار دقت و زیان آموزش و اعتبارسنجی مدل EfficientNet



ب) خطا الف) دقت

شکل (۱۲): نمودار دقت و زیان آموزش و اعتبارسنجی مدل Resnet

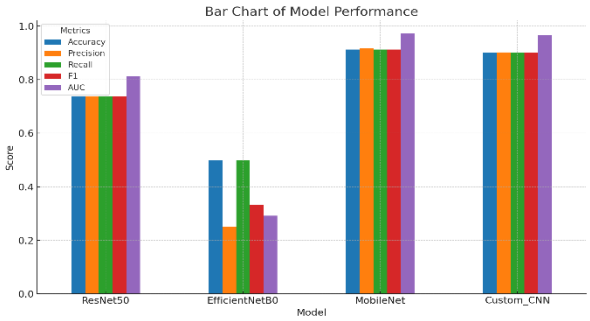


ب) خطا

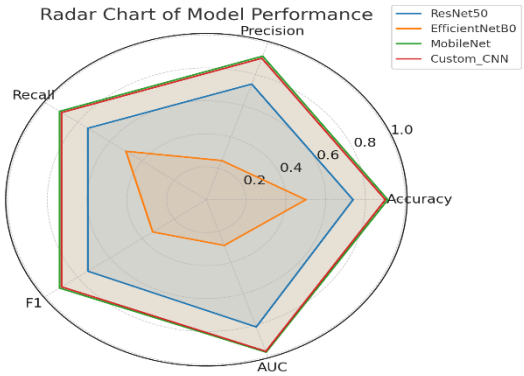
الف) دقت

شکل (۱۳): نمودار دقت و زیان آموزش و اعتبارسنجی مدل MobileNet

در نهایت، به‌منظور مقایسه جامع عملکرد مدل‌ها، نمودارهای تجمیعی و نمودار راداری[1] در شکل‌های ۱۴ و ۱۵ ترسیم شده‌اند. این نمودارها نشان می‌دهند که مدل MobileNetدر تمامی معیارهای ارزیابی (دقت، صحت، یادآوری، F1 و AUC)برتری مطلق دارد و بهترین توازن میان کارایی محاسباتی و دقت تشخیص عیوب سطحی را برقرار کرده است.



شکل (۱۴): نمودار مقایسه مدل‌های مختلف براساس معیارهای موجود



شکل (۱۵): نمودار رادار مقایسه مدل‌های مختلف

۷- نتیجه‌گیری

در این پژوهش، یک رویکرد جامع و کارآمد برای تشخیص خودکار عیب دوپوستگی در محصولات فولادی ارائه شد. سیستم پیشنهادی با بهره‌گیری از دوربین‌های صنعتیBasler ، مجموعه‌ای از تکنیک‌های پیش‌پردازش تصویر و مدل‌های شبکه‌های عصبی پیچشی آماده گردید. در مرحله پیش‌پردازش، عملیات‌هایی نظیر تشخیص لبه، برش ناحیه مورد نظر، تغییر اندازه و نرمال‌سازی به‌منظور افزایش کیفیت داده‌ها و یکنواخت‌سازی ورودی شبکه انجام گرفت. همچنین، از تکنیک داده‌افزایی برای افزایش حجم و تنوع داده‌های آموزشی و بهبود قابلیت تعمیم‌پذیری مدل استفاده شد.

نتایج تجربی نشان دادند که رویکرد پیشنهادی در مقایسه با روش‌های سنتی بازرسی، دقت، سرعت و پایداری بالاتری در تشخیص عیوب سطحی فولاد دارد. به‌ویژه، مدل MobileNetبا دقت ۹۱/۲۵ درصد و ساختاری سبک و کارایی محاسباتی بالا توانست بهترین عملکرد را در میان مدل‌های بررسی‌شده، ارائه دهد. استفاده از این سامانه در خطوط تولید می‌تواند به کاهش ضایعات، ارتقای کیفیت نهایی محصول و کاهش هزینه‌های عملیاتی در صنعت فولاد منجر شود.

در راستای توسعه آتی این پژوهش، چند مسیر تحقیقاتی جدید پیشنهاد می‌شود. نخست، اعتبارسنجی سامانه در مقیاس صنعتی از طریق پیاده‌سازی و ارزیابی عملکرد آن در محیط‌های واقعی تولید فولاد می‌تواند میزان کارایی و پایداری سامانه را در شرایط عملیاتی

[1]Radar Chart

مشخص سازد. دوم، بهره‌گیری از مدل‌های ترکیبی یادگیری عمیق با هدف افزایش دقت، پایداری و مقاومت در برابر نویز، گامی مؤثر در بهبود عملکرد مدل خواهد بود. سوم، گسترش سامانه به تشخیص چندگانه عیوب سطحی می‌تواند امکان شناسایی هم‌زمان انواع مختلف نواقص را در محصولات فولادی فراهم آورد و دامنه کاربرد سامانه را گسترش دهد. در نهایت، بهینه‌سازی مدل‌ها برای استقرار بر روی سخت‌افزارهای لبه و کاهش پیچیدگی محاسباتی به‌منظور اجرای بلادرنگ در خطوط تولید، از دیگر مسیرهای کلیدی در توسعه و صنعتی‌سازی این سامانه محسوب می‌شود.

References

[1] X. Wen, S. Jvran, H. Yu and S. Kechen, "Steel Surface Defect Recognition: A Survey," Coatings, 2023.

[2] E. Basson, "World Steel in Figures 2022," https://worldsteel.org/data/world-steel-in-figures/world-steel-in-figures-2022/, 2022.

[3] J. Qifan and C. Li, "A Survey of Surface Defect Detection of Industrial Products Based on A Small Number of Labeled Data," arXiv, 2022.

[4] Y. He, X. Wen and J. Xu, "A Semi-Supervised Inspection Approach of Textured Surface Defects under Limited Labeled Samples," Coatings, vol. 11, p. 12, 2022.

[5] Y. Chen, Y. Ding, F. Zhao, E. Zhang, Z. Wu and L. Shao, "Surface Defect Detection Methods for Industrial Products: A Review," Applied Sciences, vol. 16, p. 11, 2021.

[6] M. Chu and R. Gong, "Invariant Feature Extraction Method Based on Smoothed Local Binary Pattern for Strip Steel Surface Defect," ISIJ International, vol. 55, no. 9, pp. 1956-1962, 2015.

[7] M. Thanh Nhat Truong and S. Kim, "Automatic image thresholding using Otsu's method and entropy weighting scheme for surface defect detection," Soft Computing, vol. 22, no. 2, 2018.

[8] "Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion," Multimedia Tools and Applications, vol. 77, no. 13, pp. 16741 - 16770, 2018.

[9] S. H. Darwish and A. A. Halin, "Surface Defects Classification of Hot-Rolled Steel Strips Using Multi-directional Shearlet Features," Arabian Journal for Science and Engineering, vol. 44, p. 2925–2932, 2019.

[10] X. Liu, K. Xu, P. Zhou and D. Zhou, "Improved contourlet transform construction and its application to surface defect recognition of metals," Multidimensional Systems and Signal Processing, vol. 31, no. 11, 2020.

[11] M. Liu, Y. Liu, H. Hu and L. Nie, "Genetic Algorithm and Mathematical Morphology Based Binarization Method for Strip Steel Defect Image with Non-uniform Illumination," Journal of Visual Communication and Image Representation, vol. 37, 2016.

[12] Y.-I. Hwang, M.-K. Seo, H. G. Oh, N. Choi, G. Kim and K.-B. Kim, "Detection and Classification of Artificial Defects on Stainless Steel Plate for a Liquefied Hydrogen Storage Vessel Using Short-Time Fourier Transform of Ultrasonic Guided Waves and Linear Discriminant Analysis," Applied Sciences, vol. 12, no. 13, 2022.

[13] W. Jinjiang, F. Peilun, R. X. and Gao, "Machine vision intelligence for product defect inspection based on deep learning and Hough transform," Journal of Manufacturing Systems, vol. 51, pp. 52-60, 2019.

[14] Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Association for Computing Machinery, vol. 60, no. 6, 2017.

[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," CoRR, 2014.

[16] H. Kaiming, Z. Xiangyu, R. Shaoqing and S. Jian, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[17] S. Zhang, Q. Zhang, J. Gu and L. Su, "Visual inspection of steel surface defects based on domain adaptation and adaptive convolutional neural network," Mechanical Systems and Signal Processing, vol. 15, no. 3, 2021.

[18] H. Tajmal and S. Jongwon, "Steel Surface Defect Recognition in Smart Manufacturing Using Deep Ensemble Transfer Learning-Based Techniques," Computer Modeling in Engineering & Sciences, vol. 142, no. 1, pp. 231-250, 2025.

[19] R. Joseph, D. Santosh, G. Ross and F. Ali, "You Only Look Once: Unified, Real-Time Object Detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[20] Z. Li, X. Tian, X. Liu, Y. Liu and X. Shi, "A Two-Stage Industrial Defect Detection Framework Based on Improved-YOLOv5 and Optimized-Inception-ResnetV2 Models," Applied Sciences, vol. 12, no. 2, 2022.

[21] C. Cong, L. Hoileong and C. Ming, "Steel surface defect detection method based on improved YOLOv9," Scientific Reports, 2025.

[22] S. Youkachen, M. Ruchanurucks, T. Phatrapornnant and H. Kaneko, "Defect Segmentation of Hot-rolled Steel Strip Surface by using Convolutional Auto-Encoder and Conventional Image processing," in 2019 10th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), 2019.

[23] Y. He, K. Song, H. Dong and Y. Yan, "Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network," Optics and Lasers in Engineering, vol. 122, pp. 294-302, 2020.

[24] J. Zhang, H. Su, W. Zou, X. Gong, Z. Zhang and F. Shen, "CADN:: A weakly supervised learning-based category-aware object detection network for surface defect detection," Pattern Recogn, vol. 109, pp. 0031-3203, 2021.

[25] H. Andrew G., Z. Menglong, C. Bo, K. Dmitry, W. Weijun, W. Tobias, A. Marco and A. Hartwig, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv, 2017.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 35 - 52

Detection of Steel Bilayer Defects Using Artificial Neural Networks: A Machine Vision and Deep Learning Approach

Mostafa Majidi[1]*, Seyedamir Makinejadsanij [2]

[1] Iranian Alloy Steel Development Company, Yazd, Iran

[2] Information Technology Department, Iran Alloy Steel Company, Yazd, Iran

Article Information

Original Research Paper

Received:
2025 August 22

Accepted:
2025 November 13

Keywords:
Steel defect detection, bilayer defect, artificial neural networks, deep learning, machine vision

Corresponding Author*:
m.majidi@stu.yazd.ac.ir

Abstract

In this study, a novel approach is presented for the automatic detection of delamination defects in steel rebar using artificial neural networks and computer vision techniques. The delamination defect, which occurs due to the separation of a thin surface layer of steel during the rolling process, can lead to a reduction in mechanical properties and potential failure of the final product. Traditional inspection methods, which are typically manual and time-consuming, face significant challenges such as human error and low processing speed. In this research, a machine vision–based system was implemented immediately after the rolling process to identify delamination defects. The system employs three industrial Basler cameras positioned at 120° intervals, along with controlled xenon lighting, enabling continuous and synchronized imaging with the movement of the production line. In the initial stage, image preprocessing techniques, including normalization and segmentation, were applied to enhance image quality. Subsequently, several deep learning models were trained to classify and detect delamination defects. Evaluation of the models on real production-line data demonstrated that the MobileNet model achieved the highest accuracy of 91.25%, outperforming the other models under investigation. The obtained results indicate that the proposed system possesses industrial deployment capability and can serve as an effective and reliable alternative to traditional visual inspection methods.

: 10.22034/ABMIR.2025.23565.1159

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱٤۰٤، صفحات ۷۱ – ۵۳

بهینه‌سازی هوشمند انرژی و کنترل تطبیقی کوره‌های قوس الکتریکی در صنعت فولاد با استفاده از دوقلوی دیجیتال و شبکه‌های عصبی

سارا محمودی رشید*

مربی دانشکده مهندسی برق و کامپیوتر، دانشگاه تبریز، تبریز، ایران

مقاله پژوهشی

تاریخ دریافت:
۱٤۰٤/۰۲/۲٦

تاریخ پذیرش:
۱٤۰٤/۰۷/۰۸

کلیدواژه‌ها:
هوش مصنوعی صنعتی، کوره قوس الکتریکی، بهینه‌سازی مصرف انرژی، کنترل تطبیقی، شبکه عصبی یادگیرنده، بهره‌وری در صنعت فولاد

نویسنده‌مسئول:
s.mahmoudirashid@tabrizu.ac.ir

چکیده

در این پژوهش، یک چارچوب نوین برای بهینه‌سازی مصرف انرژی و کنترل تطبیقی عملکرد کوره‌های قوس الکتریکی در صنعت فولاد ارائه‌شده است که بر پایه تلفیق دوقلوی دیجیتال و شبکه‌های عصبی تطبیقی طراحی گردیده است. هدف اصلی، کاهش مصرف انرژی، افزایش پایداری فرآیند و ارتقاء دقت کنترل در شرایط عملیاتی متغیر است. بدین منظور، ابتدا یک مدل دوقلوی دیجیتال از فرآیند ذوب فولاد توسعه‌یافته و به‌صورت برخط به‌روزرسانی شده است. سپس، از شبکه‌های عصبی تطبیقی برای تخمین غیرخطی دینامیک فرآیند و طراحی کنترل‌کننده بهره گرفته شد. پیاده‌سازی سیستم در محیط **MATLAB/Simulink** انجام شد و عملکرد آن در سناریوهای عملیاتی گوناگون ارزیابی گردید. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی موجب کاهش متوسط ۸/۷ درصدی مصرف انرژی، بهبود ۲٤/۳ درصدی در پاسخ گذرا و افزایش ۱٤/۱۲ درصدی دقت ردیابی توان نسبت به کنترل‌کننده‌های کلاسیک شده است. این دستاوردها حاکی از اثربخشی روش ارائه‌شده در بهبود بهره‌وری انرژی و پایداری فرآیندهای حرارتی در صنعت فولاد است و می‌تواند به‌عنوان الگویی برای پیاده‌سازی سیستم‌های هوشمند مشابه در سایر فرآیندهای صنعتی در نظر گرفته شود.

: 10.22034/ABMIR.2025.23151.1128

E-ISSN: 2821-2037

۱- مقدمه

صنعت فولاد یکی از پرمصرف‌ترین بخش‌های صنعتی از نظر انرژی به‌شمار می‌رود، به‌ویژه در فرآیندهای حرارتی مانند کوره‌های قوس الکتریکی که نقش کلیدی در ذوب و بازیافت فولاد دارند[۱].

۲- پیشینه و انگیزه پژوهش

با توجه به افزایش قیمت انرژی و الزامات زیست‌محیطی، کاهش مصرف انرژی و ارتقاء بهره‌وری فرآیندهای تولیدی به یکی از اولویت‌های راهبردی شرکت‌های فولادی تبدیل‌شده است [۲]. در سال‌های اخیر، پیشرفت‌های سریع در حوزه هوش مصنوعی صنعتی و فناوری‌های دیجیتال، بستر جدیدی برای کنترل و بهینه‌سازی فرآیندهای پیچیده فراهم کرده‌اند [۳].

دوقلوی دیجیتال به‌عنوان نمایه‌ای مجازی از یک سیستم فیزیکی، امکان شبیه‌سازی دقیق، پایش بلادرنگ و به‌روزرسانی مستمر مدل‌های فرآیند را فراهم می‌سازد [٤]. از سوی دیگر، شبکه‌های عصبی تطبیقی به دلیل توانایی بالا در مدل‌سازی سیستم‌های غیرخطی و یادگیری از داده‌های جاری، ابزار قدرتمندی برای طراحی کنترل‌کننده‌های هوشمند محسوب می‌شوند [۵]. با این حال، تلفیق این دو فناوری نوظهور در محیط‌های صنعتی پیچیده مانند کوره‌های قوس الکتریکی هنوز در مراحل ابتدایی قرار دارد و نیازمند پژوهش‌های کاربردی و نظام‌مند است [٦].

در این پژوهش، با بهره‌گیری از توانمندی‌های دوقلوی دیجیتال در مدل‌سازی بلادرنگ و شبکه‌های عصبی در کنترل غیرخطی، یک چارچوب نوآورانه برای کنترل تطبیقی و بهینه‌سازی مصرف انرژی در کوره‌های قوس الکتریکی ارائه‌شده است. هدف از این تحقیق، ارائه راهکاری عملی برای افزایش بهره‌وری، کاهش هزینه‌های انرژی و تقویت پایداری در صنعت فولاد از طریق پیاده‌سازی فناوری‌های نسل چهارم صنعت است.

۳- بررسی ادبیات

کوره‌های قوس الکتریکی به‌عنوان یکی از اصلی‌ترین تجهیزات در فرآیند ذوب فولاد، نقش بسزایی در بازدهی انرژی و پایداری زیست‌محیطی صنعت فولاد ایفا می‌کنند [۷]. با این حال، ماهیت دینامیکی، غیرخطی و متغیرپذیر عملکرد کوره‌های قوس الکتریکی باعث شده است که کنترل بهینه آن به چالشی جدی در حوزه مهندسی فرآیند تبدیل شود [۸]. در دهه گذشته، پژوهش‌های متعددی به منظور بهبود عملکرد این کوره‌ها از منظر مصرف انرژی، کنترل توان ورودی و ثبات فرآیند انجام‌شده است [۹]. بسیاری از این مطالعات از الگوریتم‌های کلاسیک کنترل مانند PID[۱] یا روش‌های بهینه‌سازی خطی بهره گرفته‌اند [۱۰]؛ اما این رویکردها معمولاً در مواجهه با شرایط عملیاتی متغیر و اغتشاشات محیطی، کارایی محدودی از خود نشان داده‌اند [11].

در سال‌های اخیر، کاربرد روش‌های هوش مصنوعی به‌ویژه شبکه‌های عصبی مصنوعی در مدل‌سازی و کنترل فرآیندهای صنعتی رشد چشمگیری داشته است [۱۲]. برخی مطالعات از شبکه‌های عصبی مصنوعی برای پیش‌بینی مصرف انرژی یا طراحی کنترل‌کننده‌های پیش‌بین استفاده کرده‌اند [۱۳]. همچنین، مفهوم دوقلوی دیجیتال به‌عنوان ابزاری نوین برای شبیه‌سازی بلادرنگ و تصمیم‌گیری داده‌محور، موردتوجه صنایع پیشرفته قرارگرفته است [۱٤]. با این حال، مرور منابع موجود نشان می‌دهد که ادغام سیستماتیک دوقلوی دیجیتال با ساختارهای کنترل تطبیقی مبتنی بر یادگیری عصبی، به‌ویژه در زمینه بهینه‌سازی کوره‌های قوس الکتریکی، کمتر مورد بررسی قرارگرفته است [۱۵].

در این پژوهش، با شناسایی این خلأ، یک رویکرد ترکیبی مبتنی بر دوقلوی دیجیتال و شبکه‌های عصبی تطبیقی ارائه‌شده است تا با بهره‌گیری از ظرفیت‌های مدل‌سازی بلادرنگ و یادگیری داده‌محور، کنترل و بهینه‌سازی مصرف انرژی در کوره‌های قوس الکتریکی به‌شکلی مؤثر و پایدار تحقق یابد. مطالعات متعددی نشان داده‌اند که استفاده از شبکه‌های عصبی مصنوعی می‌تواند به بهبود دقت مدل‌سازی دینامیک فرآیندهای حرارتی مانند کوره‌های قوس الکتریکی کمک کند [۱٦]. برای مثال، پژوهش‌هایی از اوایل دهه

[۱] Proportional–integral–derivative

۲۰۱۰ به‌کارگیری ساختارهای شبکه عصبی چندلایه پرسپترون و شبکه تابع پایه شعاعی را برای شناسایی رفتار حرارتی غیرخطی در محیط‌های صنعتی پیشنهاد داده‌اند که در مقایسه با مدل‌های مبتنی بر فیزیک، دقت بالاتری داشته‌اند [۱۷].

شبکه MLPبه دلیل ساختار لایه‌های عمیق و قابلیت یادگیری روابط پیچیده، توانایی بالایی در تقریب توابع غیرخطی با دقت زیاد دارد؛ با این حال، فرآیند آموزش آن معمولاً زمان‌بر بوده و در برخی شرایط به‌ویژه در حضور داده‌های نویزی، ممکن است به کمینه‌های محلی همگرا شود. در مقابل، شبکه RBF[۱]با استفاده از توابع شعاعی به‌عنوان واحدهای مخفی، از سرعت همگرایی بالاتر و سادگی در آموزش برخوردار است و می‌تواند در تقریب نواحی محلی از داده عملکرد بسیار مطلوبی داشته باشد؛ اما نقطه ضعف اصلی آن در مدل‌سازی رفتارهای بسیار غیرخطی و گسترده نهفته است.

ترکیب این دو ساختار، موسوم به MLP-RBF[۲]، باعث می‌شود که مزایای هر دو شبکه به‌طور مکمل در یک چارچوب واحد به کار گرفته شوند. در این رویکرد، RBF نقش تقویت سرعت همگرایی و افزایش پایداری آموزش در نواحی محلی را بر عهده دارد، در حالی‌که MLP مسئولیت تعمیم‌پذیری و یادگیری الگوهای پیچیده‌تر غیرخطی در مقیاس کلی را ایفا می‌کند. نتیجه این هم‌افزایی، مدلی است که هم از دقت بالا در تقریب رفتار غیرخطی سیستم برخوردار است، هم از نظر سرعت و کارایی محاسباتی عملکرد بهتری دارد. این ویژگی‌ها به‌ویژه در محیط‌های صنعتی پویا مانند کوره‌های قوس الکتریکی، که هم رفتارهای محلی و هم دینامیک‌های پیچیده در آن‌ها وجود دارد، اهمیت زیادی پیدا می‌کند.

روش‌های کنترل تطبیقی با بهره‌گیری از الگوریتم‌های یادگیری، در سال‌های اخیر جایگزین مناسبی برای کنترل‌کننده‌های کلاسیک در شرایط متغیر شده‌اند [۱۸]. به‌ویژه، ساختارهایی مانند کنترل عصبی- تطبیقی و کنترل مبتنی بر یادگیری تقویتی در محیط‌های صنعتی غیرخطی نتایج چشمگیری داشته‌اند [۱۹]. این رویکردها امکان تنظیم پیوسته پارامترهای کنترل در شرایط کاری متغیر را فراهم کرده‌اند [۲۰].

کاربرد دوقلوی دیجیتال در صنعت فولاد یکی از زمینه‌های نوظهور در تحول دیجیتال به شمار می‌رود [۲۱]. پژوهش‌های اخیر نشان داده‌اند که ایجاد مدل مجازی هم‌زمان از کوره‌های صنعتی می‌تواند منجر به تحلیل بلادرنگ، پیش‌بینی خطا و بهینه‌سازی مصرف انرژی شود [۲۲]. گرچه هر دو حوزه دوقلوی دیجیتال و کنترل هوشمند به‌صورت جداگانه توسعه‌یافته‌اند، اما پژوهش‌های محدودی به تلفیق این دو فناوری پرداخته‌اند [۲۳]. برخی مطالعات اولیه در زمینه سیستم‌های تهویه صنعتی یا نیروگاه‌ها نشان داده‌اند که ترکیب دوقلوی دیجیتال با کنترل‌کننده‌های تطبیقی مبتنی بر شبکه عصبی مصنوعی یا منطق فازی می‌تواند اثربخشی بالایی در مدیریت بهینه مصرف انرژی داشته باشد [۲۴].

برخی پژوهش‌ها، مصرف انرژی در کوره‌های قوس الکتریکی را تا بیش از ۶۰٪ کل انرژی فرآیند فولادسازی برآورد کرده‌اند [۲۵]. تلاش‌های گسترده‌ای برای بهینه‌سازی این بخش انجام‌شده است، از جمله طراحی برنامه‌های کنترل بار هوشمند، تنظیم توالی ذوب و بهینه‌سازی نقطه کاری ولتاژ و جریان کوره [۲۶].

از دیدگاه عملی، ادغام سیستماتیک دوقلوهای دیجیتال با کنترل تطبیقی مبتنی بر شبکه عصبی در کوره‌های قوس الکتریکی با مجموعه‌ای از چالش‌های ذاتی همراه است که اهمیت و نوآوری این رویکرد را برجسته می‌سازد. نخست آنکه، کوره‌های قوس الکتریکی ماهیتی بسیار غیرخطی، پویا و وابسته به شرایط عملیاتی متغیر دارند و مدل‌سازی دقیق آن‌ها نیازمند در نظر گرفتن اثرات پیچیده‌ای مانند اغتشاشات ناشی از تغییر ترکیب قراضه و نوسانات قوس است. این مسئله سبب می‌شود که پیاده‌سازی یک دوقلوی دیجیتال دقیق و همگام‌سازی آن با واقعیت فیزیکی، فرآیندی دشوار و زمان‌بر باشد. دوم، چالش‌های جمع‌آوری داده در این محیط صنعتی قابل‌توجه است؛ چرا که سنسورها در معرض دمای بسیار بالا، میدان‌های الکترومغناطیسی شدید و شرایط سخت کاری قرار دارند که می‌تواند دقت و پایداری داده‌ها را تحت تأثیر قرار دهد. سوم، کنترل تطبیقی مبتنی بر شبکه عصبی برای آنکه بتواند عملکردی اثربخش داشته باشد، نیازمند داده‌های بلادرنگ با کیفیت

[۱] Radial basis function

[۲] Multilayer perceptron- Radial basis function

بالا و پردازش سریع است؛ در حالی‌که سرعت بالای دینامیک کوره، محدودیت‌های پردازشی سخت‌گیرانه‌ای را به همراه دارد. در نهایت، موضوع ادغام بلادرنگ میان مدل دیجیتال و سیستم کنترل نیز یک چالش اساسی است؛ زیرا هرگونه تأخیر در انتقال یا پردازش داده می‌تواند موجب کاهش اثربخشی و حتی بروز ناپایداری در سیستم شود. این موانع، در کنار پتانسیل بالای بهبود بهره‌وری انرژی و افزایش عمر تجهیزات، نشان می‌دهند که ترکیب دوقلوی دیجیتال با کنترل تطبیقی مبتنی بر شبکه عصبی در حوزه کوره‌های قوس الکتریکی نه تنها یک مسیر نوآورانه، بلکه ضرورتی برای پیشرفت آینده صنعت فولاد محسوب می‌شود.

۳-۱ شکاف‌ها و مشارکت‌های پژوهشی

با وجود پیشرفت‌های قابل‌توجه در استفاده از الگوریتم‌های هوش مصنوعی برای بهبود عملکرد فرآیندهای صنعتی، همچنان چالش‌هایی اساسی در زمینه مدل‌سازی دقیق، کنترل تطبیقی و کاهش مصرف انرژی در کوره‌های قوس الکتریکی باقی‌مانده است. بسیاری از روش‌های مرسوم کنترل، توانایی لازم برای مدیریت شرایط دینامیکی متغیر و نوسانات شدید بار در فرآیند ذوب را ندارند. همچنین، رویکردهای مبتنی بر شبکه‌های عصبی غالباً فاقد قابلیت به‌روزرسانی بلادرنگ و سازگاری با شرایط واقعی عملیاتی هستند. مشارکت‌های کلیدی این پژوهش به شرح زیر است:

- ارائه یک چارچوب یکپارچه بر پایه دوقلوی دیجیتال و شبکه‌های عصبی تطبیقی برای مدل‌سازی و کنترل فرآیند ذوب فولاد به‌صورت بلادرنگ، که قادر است شرایط متغیر و غیرخطی سیستم را با دقت بالا شبیه‌سازی و کنترل نماید.
- طراحی کنترل‌کننده تطبیقی هوشمند مبتنی بر یادگیری مستمر که توانایی سازگاری با تغییرات لحظه‌ای در فرآیند را داراست و عملکرد بهتری نسبت به روش‌های کلاسیک از خود نشان می‌دهد.
- به‌کارگیری ساختارهای ترکیبی شبکه عصبی چندلایه پرسپترون و شبکه تابع پایه شعاعی در فرآیند مدل‌سازی و کنترل، به‌گونه‌ای که مزایای دقت بالا، سرعت همگرایی مطلوب و توانایی تقریب رفتار غیرخطی سیستم به‌طور هم‌زمان محقق گردد.
- نتایج این پژوهش نه تنها نشان‌دهنده امکان‌پذیری و اثربخشی تلفیق فناوری‌های نوین مانند دوقلوی دیجیتال و یادگیری ماشین در صنعت فولاد است، بلکه الگویی عملی برای سایر صنایع انرژی‌بر در مسیر تحول دیجیتال و هوشمندسازی فراهم می‌سازد.

٤- فرمول‌بندی مسئله

در این بخش، به صورت نظام‌مند به تعریف مسئله اصلی پژوهش پرداخته می‌شود و چارچوب مفهومی آن به‌طور دقیق تشریح خواهد شد. چالش اساسی در این تحقیق، طراحی یک سامانه هوشمند برای کنترل تطبیقی و بهینه‌سازی مصرف انرژی در کوره‌های قوس الکتریکی است که قادر باشد در مواجهه با شرایط دینامیکی و متغیر فرآیند، عملکردی پایدار، دقیق و کارآمد از خود نشان دهد. با توجه به ماهیت پیچیده، غیرخطی و وابسته به زمان این نوع سیستم‌های حرارتی، بهره‌گیری از روش‌های سنتی کنترلی دیگر پاسخگوی نیازهای بهره‌وری و پایداری در صنعت فولاد نیست. از این‌رو، یک مدل دقیق از فرآیند به‌عنوان پایه طراحی کنترل‌کننده مورد نیاز است، که در این مقاله، با استفاده از مفهوم دوقلوی دیجیتال و شبکه‌های عصبی تطبیقی، این مدل توسعه‌یافته و در ادامه، راهکار پیشنهادی بر اساس آن پیاده‌سازی و ارزیابی می‌شود.

٤-١ مدل دینامیکی فرآیند ذوب در کوره قوس الکتریکی

فرآیند ذوب در کوره قوس الکتریکی به دلیل نوسانات ولتاژ، تغییر بار حرارتی و رفتار غیرخطی الکترودها، به عنوان یک سیستم پیچیده غیرخطی شناخته می‌شود. مدل دینامیکی پایه‌شده بر روی معادلات انرژی و الکتریکی قابل بیان به‌صورت (۱) است:

$$\frac{dT(t)}{dt}C = \eta \cdot P(t) - Q_{\text{loss}}(t) \qquad (۱)$$

که در آن $T(t)$ دمای لحظه‌ای فلز (C^o)، C ظرفیت گرمایی معادل سیستم (J/C^o)، η بازده انتقال حرارت، $P(t)$ توان ورودی از کوره (W) ، $Q_{loss}(t)$ تلفات حرارتی به محیط (W)است. برای مدل‌سازی تلفات، از رابطه (۲) استفاده می‌شود:

$$Q_{\text{loss}}(t) = hA\left(T(t) - T_{\text{anb}}\right) \qquad (۲)$$

که در آن h ضریب انتقال حرارت، A سطح مؤثر و T_{amb}دمای محیط است.

۴-۲ ساخت دوقلوی دیجیتال

یک دوقلوی دیجیتال به‌صورت مدل دینامیکی مجازی از فرآیند واقعی ساخته می‌شود که بتواند در زمان واقعی تطابق خود را با سیستم اصلی حفظ کند. در سیستم کوره قوس الکتریکی نمی‌توان صرفاً با داده‌های خروجی به دقت مدل اطمینان کرد، بلکه ورودی‌های فرآیند نیز باید به‌طور دقیق اندازه‌گیری و در مدل شبیه‌سازی‌شده لحاظ شوند تا تابع هزینه معنای واقعی خود را پیدا کند. در این تحقیق، برای ایجاد انطباق بین دوقلوی دیجیتال و فرآیند واقعی ورودی‌های اصلی کوره قوس الکتریکی به صورت بردار ورودی (۳) در نظر گرفته‌شده‌اند:

$$u(i) = \begin{bmatrix} P_{el}(i) \\ m_{\text{scrap}}(i) \\ m_{\text{oxy}}(i) \\ m_{\text{coal}}(i) \\ Q_{\text{cool}}(i) \end{bmatrix} \quad (۳)$$

که در آن $P_{el}(i)$توان الکتریکی تزریق‌شده به کوره در لحظه i بر حسب (MW)، $m_{\text{scrap}}(i)$ جرم شارژ قراضه فلزی در هر سیکل (بر حسب تن)، $m_{\text{oxy}}(i)$ میزان تزریق اکسیژن (Nm^3/min)، $m_{\text{coal}}(i)$، $Q_{\text{cool}}(i)$، میزان تزریق کربن جانبی (kg/min) میزان خنک‌کاری آب یا سیال در دیواره‌ها (kW) یا (kJ/min).

مدل دوقلوی دیجیتال بر اساس یک معادله موازنه انرژی دینامیکی به‌صورت (۴) بیان‌شده است:

$$\begin{aligned} & C_{\text{eq}} \frac{dT_{\text{pred}}(t)}{dt} \\ & = \eta_{\text{el}} P_{\text{el}}(t) + \eta_{\text{oxy}} H_{\text{oxy}} \left(m_{\text{oxya}}(t)\right) \\ & + \eta_{\text{coal}} H_{\text{coal}} \left(m_{\text{coal}}(t)\right) \\ & - Q_{\text{loss}} \left(T_{\text{pred}}(t), Q_{\text{cool}}(t)\right) \end{aligned} \quad (۴)$$

که در آن C_{eq} ظرفیت گرمایی مؤثر بار در کوره، $\eta_{\text{el}}, \eta_{\text{oxy}}, \eta_{\text{coal}}$ ضرایب بازدهی ورودی‌های الکتریکی، اکسیژن و کربن، H_{oxy}و H_{coal} انرژی شیمیایی آزادشده از واکنش‌های اکسیژن و کربن، Q_{loss} اتلاف حرارتی وابسته به دمای بار و نرخ خنک‌کاری. دمای پیش‌بینی‌شده مدل دیجیتال در گام i برابر است با رابطه (۵):

$$T_{\text{pred}}(i;\theta) = f(u(i),\theta) \quad (۵)$$

که در آن $u(i)$ همان بردار ورودی‌های واقعی است که توسط حسگرها در فرآیند عملیاتی ثبت‌شده و θپارامترهای قابل تنظیم مدل (مانند بازدهی‌ها و ضرایب انتقال حرارت) است.

به‌منظور انطباق مدل با داده‌های واقعی، یک فیلتر تطبیقی برخط به‌کار گرفته می‌شود که با کمینه‌سازی تابع خطای پیش‌بینی، پارامترهای مدل را توسط معادله (۶) به‌روزرسانی می‌کند:

$$J(\theta) = \sum_{i=1}^{N} \left(T_{\text{meas}}(i) - T_{\text{pred}}(i;\theta)\right)^2 \quad (۶)$$

که در آن θپارامترهای بهینه مدل، T_{meas}دمای واقعی اندازه‌گیری‌شده، T_{pred}خروجی مدل پیش‌بینی‌شده توسط دوقلوی دیجیتال است. در این حالت، الگوریتم بهینه‌سازی (مانند گرادیان کاهشی یا فیلتر تطبیقی کالمن) پارامترهای θ را به‌گونه‌ای تنظیم می‌کند که اختلاف بین دمای واقعی و دمای شبیه‌سازی‌شده حداقل شود.

۴-۳ طراحی شبکه عصبی تطبیقی برای مدل‌سازی معکوس فرآیند

برای تخمین معکوس رفتار غیرخطی سیستم و پیش‌بینی توان ورودی به‌ازای دمای هدف، از شبکه عصبی ترکیبی MLP–RBF شبکه عصبی چندلایه پرسپترون و شبکه تابع پایه شعاعی معادله (۷) استفاده می‌شود:

$$P_{\text{est}}(t) = f_{MLP}\left(T(t), \frac{dT(t)}{dt}, T_{\text{ref}}(t)\right) + f_{RBF}(T(t)) \quad (۷)$$

که f_{MLP}بخش یادگیرنده با ساختار پرسپترون چندلایه، f_{RBF} تابع تقریب موضعی با گره‌های شعاعی، $T_{ref}(t)$دمای مرجع هدف‌گذاری‌شده است. ترکیب این دو ساختار، هم دقت بالا و هم توانایی تطبیق سریع در شرایط متغیر را تضمین می‌کند.

۴-۴ طراحی کنترل‌کننده تطبیقی مبتنی بر مدل شبکه عصبی

بر اساس مدل به‌دست‌آمده، کنترل‌کننده‌ای طراحی می‌شود که خروجی مطلوب T_{ref}را با کمترین مصرف انرژی دنبال کند. سیگنال کنترلی به‌صورت تطبیقی (۸) به‌روزرسانی می‌شود:

$$u(t) = \alpha u(t-1) + \beta \cdot \left(T_{\text{ref}}(t) - T(t)\right) + \gamma \cdot \frac{d}{dt}\left(T_{\text{ref}}(t) - T(t)\right) \quad (۸)$$

که ضرایب α، β و γ به‌صورت برخط و با استفاده از گرادیان کاهش خطا به‌روزرسانی می‌شوند:

$$\theta(t+1) = \theta(t) - \eta \cdot \nabla_\theta J(\theta) \quad (۹)$$

٤-٤-١ به‌روزرسانی برخط

منظور از به‌روزرسانی برخط آن است که ضرایب کنترلی رابطه (٤) در هر گام نمونه‌برداری، همزمان با ورود داده‌های جدید اندازه‌گیری‌شده از کوره، بازتنظیم می‌شوند؛ یعنی بدون توقف فرایند و بر مبنای داده‌های جاری رابطه (۱۰):

$$\begin{aligned} u(t) &= \alpha(t)u(t-1) + \beta(t)e(t) \\ &\quad + \gamma(t)\dot{e}(t), \\ e(t) &= T_{\text{ref}}(t) - T(t) \end{aligned} \quad (۱۰)$$

در این‌جا $\dot{e}(t)$ مشتق خطا است که با یک مشتق‌گیر فیلترشده توسط معادله (۱۱) محاسبه می‌شود:

$$\begin{aligned} \dot{e}(t) &= (1-\lambda)\dot{e}(t-1) \\ &\quad + \lambda \frac{e(t) - e(t-1)}{\Delta t}, \\ 0 < \lambda &< 1 \end{aligned} \quad (۱۱)$$

٤-٤-٢ گردآوری و همگام‌سازی داده

داده‌ها در افق زمانی $\{t_k\}_{k=1}^{N}$ با دوره نمونه‌برداری Δt جمع‌آوری می‌شوند:

- دما $T(t_k)$: از پایرومتر نوری/ترموکوپل‌های تماس‌پذیر (رزولوشن 1°C و خطای کالیبراسیون ±0.5%).
- توان لحظه‌ای $P(t_k)$ (خروجی مبدل/ترانس): از پاورآنالایزر (رزولوشن 0/1 MW).
- دبی اکسیژن/مواد افزودی :از فلومترهای جرمی (برای دوقلو دیجیتال و اعتبارسنجی).

برای حذف نویز اندازه‌گیری، از میانگین متحرک درجه ۳ مطابق رابطه (۱۲) استفاده می‌کنیم:

$$\tilde{T}(t_k) = \frac{1}{3}\left(T(t_k) + T(t_{k-1}) + T(t_{k-2})\right) \quad (۱۲)$$

و سپس $\tilde{e}(t_k) = T_{\text{ref}}(t_k) - \tilde{T}(t_k)$ مبنای قانون تطبیق قرار می‌گیرد. برچسب زمانی یکسان موجب همگام‌سازی کانال‌ها می‌شود تا $e, \dot{e}, u$ در یک t قابل اتکا باشند.

٤-٤-٣ قید وابستگی ضرایب (کاهش درجه آزادی)

برای جلوگیری از عدم‌تشخیص ضرایب، قیود (۱۳) اعمال می‌شود:

$$\begin{aligned} &\alpha(t) \geq 0, \beta(t) \geq 0, \gamma(t) \geq 0, \\ &\alpha(t) + \beta(t) + \gamma(t) = 1 \end{aligned} \quad (۱۳)$$

بدین‌ترتیب فقط دو پارامتر مستقل تنظیم می‌شوند و سومی از قید جمع به‌دست می‌آید (آنالوگِ «یک منهای ضریب» در حالت دوعبارتی).

٥-٤ تعریف تابع هدف چندمعیاره

هدف کنترل‌کننده نه تنها رسیدن به دمای مطلوب بلکه بهینه‌سازی مصرف انرژی و حفظ پایداری حرارتی در برابر اغتشاشات محیطی است. بنابراین، یک تابع هدف چندمعیاره به صورت (۱٤) تعریف می‌شود:

$$J(\theta) = w_1 \cdot e_T^2(t) + w_2 \cdot \Delta u^2(t) + w_3 \cdot \left(\frac{dT(t)}{dt}\right)^2 \quad (۱٤)$$

که در آن $\boldsymbol{e(t) = T_{ref}(t) - T(t)}$ خطای دمایی، $\boldsymbol{\Delta u(t) = u(t) - u(t-1)}$ تغییرات ناگهانی در سیگنال کنترلی، $\frac{dT(t)}{dt}$ نرخ تغییرات دما به‌عنوان معیار پایداری و ضرایب w_1 ، w_2 و w_3 وزن‌دهی اهداف مختلف سیستم هستند.

ضرایب w_1، w_2 و w_3 نمی‌توانند کاملاً مستقل از یکدیگر تنظیم شوند، زیرا این امر منجر به بیش‌پارامتری می‌شود. برای رفع این مشکل، قید (۱۵) اعمال می‌شود:

$$w_1 \geq 0, w_2 \geq 0, w_3 \geq 0, w_1 + w_2 + w_3 = 1 \quad (۱۵)$$

در نتیجه، تنها دو ضریب مستقل باقی می‌مانند و ضریب سوم به‌طور خودکار از رابطه بالا محاسبه می‌شود. برای پیاده‌سازی عملی، دو روش متداول وجود دارد:

پارامترسازی **Softmax** توسط (۱٦):

$$\begin{aligned} w_1 &= \frac{e^a}{e^a + e^b + e^c}, \\ w_2 &= \frac{e^b}{e^a + e^b + e^c}, w_3 = \frac{e^c}{e^a + e^b + e^c}, \end{aligned} \quad (۱٦)$$

که در آن $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) \in \mathbb{R}$ متغیرهای بدون قید هستند و ضرایب همیشه نرمال و نامنفی باقی می‌مانند.

- پروژکشن روی سیمپلکس:

پس از هر گام به‌روزرسانی ضرایب با روش گرادیان، بردار $\theta = [w_1, w_2, w_3]^T$ **روی مجموعه** $S = \{\theta \in \mathbb{R}^3 \mid \theta \geq 0, 1^T\theta = 1\}$ **پرتاب می‌شود.**

در اصل، مسأله کنترلی ماهیت چندهدفه دارد، زیرا کمینه‌سازی خطای دما، کاهش تلاش کنترلی و تضمین پایداری حرارتی اهدافی متعارض هستند. برای حل عملی، از روش اسکالاری‌سازی استفاده می‌شود که مسأله چندهدفه را به یک تابع تک‌هدفه وزن‌دار تبدیل می‌کند. این کار امکان استفاده از روش‌های کلاسیک کنترل تطبیقی و بهینه‌سازی مانند گرادیان کاهشی و کنترل پیش‌بین مدل را فراهم می‌سازد.

ضرایب w_1، w_2 و w_3 به‌صورت تجربی یا با روش‌های بهینه‌سازی (مانند الگوریتم ژنتیک) تعیین می‌شوند. به طور مشخص، در الگوریتم ژنتیک یک تابع هدف تعریف می‌شود که میزان خطا یا اختلاف بین مقادیر واقعی (به‌دست‌آمده از اندازه‌گیری یا شبیه‌سازی) و مقادیر تخمینی (محاسبه‌شده توسط مدل پیشنهادی) را ارزیابی می‌کند. هدف اصلی این است که ضرایب به‌گونه‌ای انتخاب شوند که این اختلاف حداقل شود. تابع برازش مورد استفاده به شکل (۱۷) تعریف شد:

$$J = \sum_{k=1}^{N} \left(y_{\text{measured}}(k) - y_{\text{estimated}}(k; \alpha_1, \alpha_2, \alpha_3)\right)^2 \quad (۱۷)$$

که در آن:

- $y_{\text{measured}}(k)$ داده واقعی یا مرجع در لحظه k است،
- $y_{\text{estimated}}(k; \alpha_1, \alpha_2, \alpha_3)$ خروجی مدل پیشنهادی با ضرایب $\alpha_1, \alpha_2, \alpha_3$،
- N تعداد کل نمونه‌های داده،
- $\alpha_1, \alpha_2, \alpha_3$ ضرایب وزنی مورد نظر.

با توجه به محدودیت وابستگی بین ضرایب، شرط (۱۸) نیز در فرآیند بهینه‌سازی اعمال می‌شود:

$$\alpha_1 + \alpha_2 + \alpha_3 = 1, \alpha_i \geq 0 \quad (۱۸)$$

بنابراین، عملاً تنها دو ضریب مستقل هستند و ضریب سوم به‌صورت خودکار از رابطه بالا تعیین می‌شود. الگوریتم ژنتیک با ایجاد جمعیتی اولیه از ضرایب، محاسبه مقدار تابع J و استفاده از عملگرهای انتخاب، جهش و تقاطع، به‌صورت تکراری مقادیر بهینه ضرایب را پیدا می‌کند. معیار توقف الگوریتم رسیدن به حداقل خطا یا تعداد مشخصی نسل تکرار است.

۴-۶ مدل بهبود یافته تخمین نرخ تغییرات انرژی در فرآیند

برای در نظر گرفتن اثر دینامیک انتقال انرژی درون کوره، مدل انرژی بهبودیافته‌ای به‌کار می‌رود که نرخ تغییر انرژی تجمعی را توسط (۱۹) شامل می‌شود:

$$\frac{dE_{\text{totd}}(t)}{dt} = \eta \cdot P(t) - \left(hA\left(T(t) - T_{\text{anb}}\right) + R \cdot I^2(t)\right) \quad (۱۹)$$

در اینجا $E_{total}(t)$ انرژی تجمع‌یافته در فرآیند ذوب، R مقاومت الکتریکی مسیر جریان، $I(t)$ جریان ورودی لحظه‌ای به الکترودها است. این مدل اجازه می‌دهد رفتار حرارتی سیستم با دقت بیشتری تخمین زده‌شده و برای پیش‌بینی دقیق‌تر مصرف انرژی استفاده شود.

۴-۷ به‌روزرسانی تطبیقی ضرایب کنترل‌کننده با استفاده از قانون یادگیری

برای بهبود انعطاف‌پذیری در برابر تغییرات فرآیندی، ضرایب کنترل‌کننده به‌صورت تطبیقی و با استفاده از قانون یادگیری مبتنی بر گرادیان خطا (۲۰) اصلاح می‌شوند:

$$\frac{d\theta_i(t)}{dt} = -\lambda \cdot \frac{\partial J(t)}{\partial \theta_i(t)} \quad (۲۰)$$

که در آن θ_i شامل ضرایب α، β و γ کنترل‌کننده است. λ نرخ یادگیری و $\frac{\partial J(t)}{\partial \theta_i(t)}$ مشتق تابع هزینه نسبت به هر پارامتر است.

۴-۸ معیار پایداری با استفاده از تابع لیاپانوف

برای تضمین پایداری سیستم در برابر اغتشاشات و تغییرات مدل، تابع لیاپانوف (۲۱) تعریف می‌شود:

$$V(t) = \frac{1}{2}e_T^2(t) + \frac{1}{2}\sum_i (\theta_i(t) - \theta_i^*)^2 \quad (۲۱)$$

و با اطمینان از اینکه مشتق آن منفی مطابق (۲۲) باشد:

$$\frac{dV(t)}{dt} \le 0 \quad (۲۲)$$

می‌توان نشان داد که سیستم به‌صورت مجانبی پایدار است و در صورت وجود اغتشاش، پاسخ از نوسان بیش‌ازحد محافظت می‌شود.

۴-۹ ارتباط دوقلوی دیجیتال با کنترلر در حلقه بسته

دوقلوی دیجیتال نه‌تنها برای شبیه‌سازی، بلکه در حلقه بسته برای تصحیح خطای مدل و ارسال پیش‌بینی‌های اصلاح‌شده به کنترل‌کننده مورد استفاده (۲۳) قرار می‌گیرد:

$$u_{\text{corr}}(t) = u_{NN}(t) + \kappa \cdot \left(T_{\text{meas}}(t) - T_{DT}(t)\right) \quad (۲۳)$$

که در آن $\boldsymbol{u_{NN}(t)}$ خروجی اولیه شبکه عصبی، $\boldsymbol{T_{DT}(t)}$ دمای تخمین‌زده‌شده توسط دوقلوی دیجیتال و $\boldsymbol{\kappa}$ ضریب تطبیق برای اصلاح کنترل است.

۴-۱۰ ساختار شبکه عصبی تطبیقی

در این پژوهش، یک شبکه عصبی چندلایه با توانایی یادگیری برخط به‌منظور تقریب دینامیک ناشناخته سیستم در نظر گرفته‌شده است. ساختار کلی شبکه به‌صورت (۲۴) تعریف می‌شود:

$$y_{NN}(t) = \sum_{i=1}^{n} w_i(t) \cdot \sigma\left(\sum_{j=1}^{m} v_{ij}(t) \cdot x_j(t) \cdot b_i\right) \quad (۲۴)$$

که در آن $\boldsymbol{x_j(t)}$ ورودی‌های شبکه (شامل دما، نرخ تغییر دما، جریان، توان لحظه‌ای)، $\boldsymbol{\sigma(.)}$ تابع فعال‌سازی، به‌صورت سیگموید، $\boldsymbol{w_i(t)}$ و $\boldsymbol{v_{ij}(t)}$ وزن‌های ورودی و خروجی شبکه که به‌صورت تطبیقی به‌روزرسانی می‌شوند. همچنین $\boldsymbol{b_i}$ بایاس لایه میانی است.

۴-۱۱ قانون یادگیری اصلاح‌شده مبتنی بر خطای مدل

برای بهبود تطابق شبکه با تغییرات دینامیکی فرآیند، یک قانون یادگیری اصلاح‌شده بر اساس خطای مدل بین دوقلوی دیجیتال و واقعیت طراحی‌شده (۲۵) است:

$$\Delta w_i = -\eta \cdot e_T(t) \cdot \sigma_i(t) + \xi \cdot \left(T_{\text{real}}(t) - T_{DT}(t)\right) \quad (۲۵)$$

که در آن $\boldsymbol{\eta}$ نرخ یادگیری اصلی، $\boldsymbol{\xi}$ نرخ اصلاح ناشی از عدم تطابق بین دوقلوی دیجیتال و فرآیند واقعی است. این اصلاح موجب همگرایی سریع‌تر و دقت بالاتر شبکه در تخمین مدل می‌شود.

۴-۱۲ طراحی کنترل‌کننده بر پایه شبکه عصبی پیش‌بینی‌کننده

خروجی کنترل به‌صورت پیش‌بینی‌شده برای چند مرحله آینده (۲۶) محاسبه می‌شود:

$$u(t) = \arg\min_{u} \sum_{K=1}^{N_P} \left(T_{\text{ref}}(t+k) - T_{NN}(t+k \mid t)\right)^2 + \lambda \cdot \sum_{k=1}^{N_c} (\Delta u(t+k-1))^2 \quad (۲۶)$$

که در آن $\boldsymbol{N_p}$ افق پیش‌بینی، $\boldsymbol{N_c}$ افق کنترل، $\boldsymbol{T_{NN}(t+k|t)}$ دمای پیش‌بینی‌شده توسط شبکه عصبی و $\boldsymbol{\Delta u(t)}$ تغییرات سیگنال کنترلی است. این ساختار مشابه کنترل پیش‌بین مدل عمل می‌کند اما با از شبکه عصبی تطبیقی به‌جای مدل ریاضی صریح بهره‌گیری می‌کند.

۴-۱۳ ترکیب کنترل‌کننده با دوقلوی دیجیتال به‌منظور بهبود قابلیت اطمینان

جهت افزایش اعتمادپذیری در تصمیمات کنترلی، دوقلوی دیجیتال به‌عنوان بلوک بازخورد برای بررسی صحت تصمیمات کنترلی وارد حلقه (۲۷) می‌شود:

$$\text{If } \left|T_{DT}(t+1 \mid u(t)) - T_{\text{pred}}(t+1)\right| > \delta, \text{ then } u(t) \leftarrow u(t) - \epsilon \cdot \text{sign}(e_T(t)) \quad (۲۷)$$

در این رابطه δ آستانه مجاز اختلاف پیش‌بینی، ϵ ضریب تنظیم تدریجی کنترلی و $\boldsymbol{T_{pred}}$ خروجی پیش‌بینی اولیه کنترل‌کننده است. این مکانیسم به تصمیم‌گیرنده اجازه می‌دهد تا کنترل را با توجه به بازخورد دوقلوی دیجیتال تصحیح کرده و از اشتباهات ناشی از مدل‌سازی جلوگیری کند.

۴-۱۴ تحلیل پیچیدگی محاسباتی

در ارزیابی روش پیشنهادی، پیچیدگی محاسباتی هر مرحله تحلیل و با روش‌های مقایسه‌ای سنجیده شد. الگوریتم پیشنهادی شامل

سه بخش اصلی است:۱- مدل‌سازی دینامیکی و استخراج ویژگی‌ها، ۲- تخمین پارامترها با استفاده از فیلتر/بهینه‌ساز معرفی‌شده و ۳- یادگیری تطبیقی بر پایه شبکه عصبی/ماشین یادگیرنده.

- در بخش مدل‌سازی، محاسبات خطی ماتریسی با مرتبهٔ $O(n^2)$ انجام می‌شود که با توجه به ابعاد متوسط سیستم‌های صنعتی قابل مدیریت است.
- بخش تخمین پارامترها عمدتاً بر اساس روش‌های بازگشتی و فیلترهای حالت عمل می‌کند که پیچیدگی آن حدود $O(n^3)$است؛ اما با توجه به ساختار تکرارشونده، این هزینه در اجراهای آنلاین به شدت کاهش می‌یابد و در عمل نزدیک به $O(n^2)$ باقی می‌ماند.
- در بخش یادگیری تطبیقی، با استفاده از شبکه‌های عصبی سبک‌وزن، پیچیدگی یادگیری در مرحلهٔ آموزش برابر با $O(N \cdot d)$که در آن N تعداد نمونه‌ها و d تعداد ویژگی‌ها است) بوده و در مرحلهٔ استنتاج تنها $O(d)$ خواهد بود. این موضوع اجرای بلادرنگ را تضمین می‌کند.

به‌طورکلی، پیچیدگی محاسباتی روش پیشنهادی در مقایسه با روش‌های مرجع هم‌رده یا پایین‌تر است، در حالی که دقت بالاتری را فراهم می‌آورد. این توازن میان دقت و هزینه محاسباتی موجب می‌شود که الگوریتم قابلیت پیاده‌سازی عملی در سیستم‌های صنعتی مقیاس‌متوسط و بزرگ را داشته باشد. به‌ویژه در محیط‌های دارای منابع سخت‌افزاری محدود، روش پیشنهادی به دلیل ساختار بازگشتی و طراحی ماژولار، امکان استقرار آسان را فراهم می‌سازد. تحلیل پیچیدگی محاسباتی نشان می‌دهد که روش پیشنهادی با وجود افزایش دقت، از نظر هزینه محاسباتی مقیاس‌پذیر بوده و برای کاربردهای بلادرنگ در سیستم‌های صنعتی مناسب است.

به‌منظور روشن‌تر ساختن ابعاد مقیاس‌پذیری و عملی بودن روش پیشنهادی، یک مقایسهٔ دقیق از پیچیدگی محاسباتی آن با برخی روش‌های مرسوم در ادبیات شامل RLS[۱]، EKF[۲]و شبکه عصبی ساده در جدول (۱) ارائه گردید. این مقایسه به ما اجازه می‌دهد تا نشان دهیم که اگرچه روش پیشنهادی از لحاظ ترکیب چندین ماژول (مانند فیلتر، الگوریتم‌های بهینه‌سازی و ساختار تطبیقی) پیچیدگی بیشتری دارد، اما به‌دلیل طراحی بهینه و بهره‌گیری از ساختارهای ماتریسی ساده، نرخ رشد پیچیدگی آن همچنان در حد چندجمله‌ای باقی می‌ماند و لذا در سیستم‌های با ابعاد بزرگ و اجرای بلادرنگ نیز عملی خواهد بود.

جدول (۱): مقایسه پیچیدگی محاسباتی روش پیشنهادی با سایر روش‌ها

روش	نوع عملیات غالب	مرتبه پیچیدگی محاسباتی	توضیح
RLS	به‌روزرسانی ماتریس معکوس	$O(n^2)$	کارایی بالا، ولی حساس به نویز و خطاهای گرد کردن
EKF	خطی‌سازی و به‌روزرسانی کوواریانس	$O(n^3)$	کارایی بالا، ولی حساس به نویز و خطاهای گرد کردن
شبکه عصبی ساده	ضرب ماتریس و توابع فعال‌سازی	$O(N \cdot d)$	ساده‌تر ولی نیازمند داده‌های زیاد برای آموزش
روش پیشنهادی	ترکیب بهینه‌سازی ماتریسی و یادگیری تطبیقی	$O(n^2 \log n)$	ساده‌تر ولی نیازمند داده‌های زیاد برای آموزش

۵- نتایج و بحث

در این بخش، نتایج شبیه‌سازی ارائه‌شده و عملکرد روش پیشنهادی از جنبه‌های مختلف مورد تحلیل و بررسی قرار می‌گیرد. هدف از این تحلیل، ارزیابی دقت، پایداری و کارایی الگوریتم طراحی‌شده در شرایط مختلف عملیاتی و مقایسه آن با روش‌های موجود است. همچنین، تأثیر به‌کارگیری شبکه عصبی تطبیقی، دوقلوی دیجیتال و کنترل پیش‌بین در بهبود پاسخ سیستم و افزایش قابلیت اطمینان کنترل بررسی می‌شود. نمودارها، جداول و مقایسه‌های عددی به‌منظور تبیین بهتر نتایج به کار رفته‌اند و بحث‌های ارائه‌شده بر

[۱] Recursive Least Squares

[۲] Extended Kalman Filter

مبنای داده‌های حاصل از شبیه‌سازی صورت گرفته‌اند تا تصویری دقیق از عملکرد واقعی سیستم ارائه شود. در ادامه یک مثال صنعتی برای پیاده‌سازی روش ارائه‌شده در مقاله، شامل بهینه‌سازی هوشمند انرژی و کنترل تطبیقی در کوره‌های قوس الکتریکی در صنعت فولاد ارائه می‌شود.

۵-۱ مثال صنعتی: کنترل تطبیقی و بهینه‌سازی مصرف انرژی در کوره قوس الکتریکی با کمک دوقلوی دیجیتال و شبکه عصبی

کوره قوس الکتریکی یکی از تجهیزات کلیدی در صنعت فولاد است که با اعمال جریان الکتریکی شدید، قراضه فلزی را ذوب می‌کند. مصرف انرژی در این فرآیند بسیار بالا است و پارامترهایی مانند دما، نرخ ذوب، مقاومت الکترودها و نرخ ورود مواد تأثیر زیادی بر عملکرد دارند. در این مثال، هدف پیاده‌سازی یک سیستم کنترل هوشمند بر پایه شبکه عصبی تطبیقی و دوقلوی دیجیتال است که بتواند مصرف انرژی را بهینه کرده و عملکرد کوره را پایدار و قابل‌اعتماد نگه دارد. برای مدل‌سازی دینامیک دما درون کوره، از معادله انرژی استفاده می‌کنیم. برای تخمین دقیق مدل ناشناخته کوره، از یک شبکه عصبی استفاده می‌شود. وزن‌ها از قانون یادگیری تطبیقی با بازخورد دوقلوی دیجیتال به‌روزرسانی می‌شوند. کنترل‌کننده توان ورودی به الکترودها را به‌صورت پیش‌بینانه بهینه می‌کند. این مثال نشان می‌دهد چگونه می‌توان با ترکیب شبکه عصبی تطبیقی، دوقلوی دیجیتال و کنترل پیش‌بین، مصرف انرژی کوره را کاهش داده و پایداری فرآیند ذوب را افزایش داد. این رویکرد قابلیت تعمیم به دیگر فرآیندهای صنعتی پرانرژی را نیز دارد. در ادامه، در جدول (۲) پارامترهای عددی مورد استفاده برای شبیه‌سازی فرآیند کنترل تطبیقی و بهینه‌سازی انرژی در کوره قوس الکتریکی ارائه می‌شود. این مقادیر به‌صورت تقریبی و بر پایه داده‌های صنعتی مرسوم انتخاب شده‌اند تا شرایط یک سیستم واقعی را شبیه‌سازی کنند. هدف، بررسی عملکرد کنترل‌کننده پیش‌بین عصبی در تعامل با دوقلوی دیجیتال برای رسیدن به پایداری دما و کاهش انرژی مصرفی است.

فرض کنید در یک بازه زمانی نمونه‌برداری داریم:

- توان الکتریکی تزریقی $P_{\text{in}}(i) = 30\text{MW}$،
- جرم شارژ شده قراضه $M_{\text{in}}(i) = 5\text{ton}$،
- نرخ تزریق اکسیژن $Q_{O_2}(i) = 1000\text{Nm}^3/\text{h}$.

جدول (۲): پارامترهای عددی برای کنترل و مصرف انرژی

پارامتر	نماد	مقدار عددی	واحد
ظرفیت گرمایی معادل	C	12/000	J/°C
توان نامی ورودی	P_{max}	50/000	kW
نرخ تلفات حرارتی پایه	$Q_{loss,0}$	5/000	kW
دمای مرجع ذوب	T_{ref}	1/650	°C
نرخ یادگیری شبکه عصبی	η	0/001	—
نرخ اصلاح از دوقلو	ξ	0/002	—
افق پیش‌بینی	N_p	10	—
افق کنترل	N_c	5	—
ضریب وزن‌دهی تغییرات کنترل	λ	0/05	—
آستانه اختلاف	δ	15	°C
نرخ اصلاح کنترلی	ε	0/5	—

مدل دما در این حالت به‌صورت ساده‌شده به شکل (۲۸) نوشته می‌شود:

$$T_{\text{pred}}(i;\theta) = \theta_1 P_{\text{in}}(i) - \theta_2 M_{\text{in}}(i) + \theta_3 Q_{O_2}(i) \quad (۲۸)$$

با مقادیر اولیه (۲۹):

$$\theta_1 = 0/12°\text{C/MW}, \theta_2 = 8°\text{C/ ton}, \theta_3 = 0/05°\text{C/(Nm}^3\text{/h)}. \quad (۲۹)$$

بنابراین با جایگذاری مقادیر عددی (۳۰):

$$\begin{aligned} T_{\text{pred}}(i;\theta) &= 0/12 \times 30 - 8 \times 5 + 0/05 \\ &\times 1000 \\ &= 3/6 - 40 + 50 \\ &= 13/6°\text{C} \end{aligned} \quad (۳۰)$$

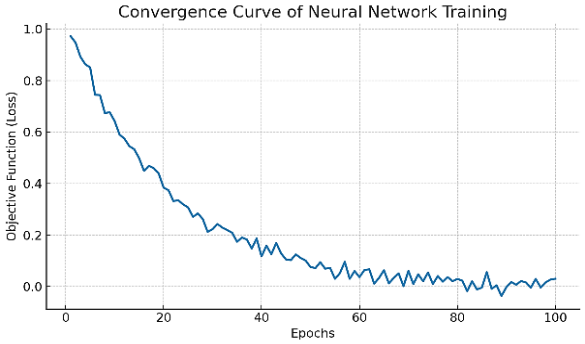
اگر دمای واقعی اندازه‌گیری‌شده در همان لحظه $T_{\text{meas}}(i) =$ $15°\text{C}$ باشد، آنگاه خطای مربعی (۳۱):

$$\begin{aligned} e(i) &= \left(T_{\text{meas}}(i) - T_{\text{pred}}(i;\theta)\right)^2 \\ &= (15 - 13/6)^2 = 1/96 \end{aligned} \quad (۳۱)$$

این خطا در تابع هزینه لحاظ شده و با استفاده از الگوریتم‌های بهینه‌سازی، مثلاً RLS یا گرادیان کاهشی، پارامترهای θ_1، θ_2 و θ_3 به‌صورت برخط اصلاح می‌شوند تا انطباق بهتری بین دوقلوی دیجیتال و فرآیند واقعی حاصل شود. بدین ترتیب، ورودی‌های اصلی فرآیند همواره اندازه‌گیری شده و در مدل لحاظ می‌شوند و

این مسئله تضمین می‌کند که خروجی‌های پیش‌بینی‌شده قابلیت مقایسه واقعی با داده‌های صنعتی داشته باشند.

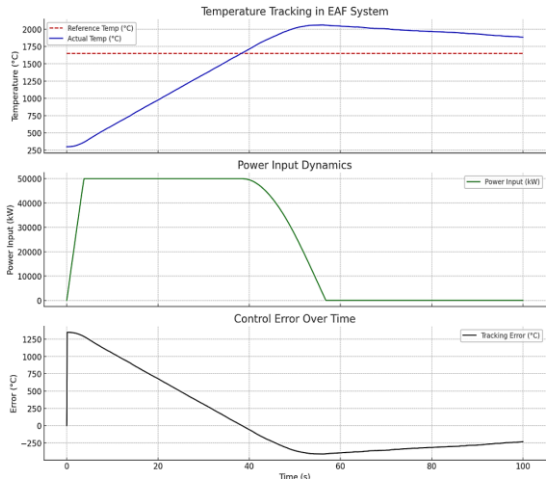
منحنی تغییرات تابع هدف در طول آموزش شبکه عصبی بیانگر فرایند همگرایی مدل در شکل (۱) است. در ابتدای آموزش، مقدار تابع هدف نسبتاً بزرگ است زیرا وزن‌ها و بایاس‌های شبکه به‌صورت تصادفی مقداردهی اولیه شده‌اند. با تکرار چرخه‌های آموزشی و اعمال الگوریتم بهینه‌سازی، خطا به‌تدریج کاهش‌یافته و شبکه به سمت مقادیر بهینه حرکت می‌کند. روند نزولی منحنی نشان می‌دهد که مدل به‌طور مؤثر الگوهای داده را یاد گرفته و توانسته است خطا را کاهش دهد.



شکل (۱): منحنی تغییرات تابع هدف در حین آموزش شبکه عصبی

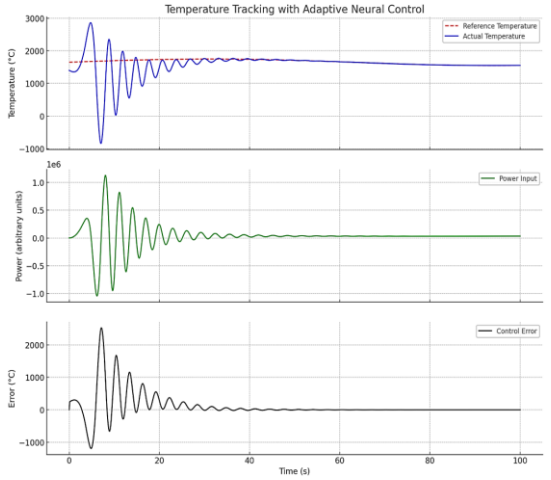
در مراحل پایانی آموزش، تغییرات تابع هدف بسیار ناچیز شده و منحنی به یک مقدار تقریباً ثابت همگرا می‌شود که نشان‌دهنده رسیدن شبکه به پایداری نسبی و یادگیری کافی است. این رفتار صحت عملکرد الگوریتم بهینه‌سازی و کفایت ساختار انتخاب‌شده برای شبکه را تأیید می‌کند.

در شکل (۲) عملکرد یک کنترل‌کننده الهام‌گرفته از شبکه عصبی برای تنظیم دمای کوره قوس الکتریکی مورد بررسی قرارگرفته است. دمای مرجع ثابت در نظر گرفته‌شده و کنترل‌کننده تلاش می‌کند با توجه به خطای دما و نویز تصادفی، سیگنال توان ورودی را به‌گونه‌ای تنظیم کند که دمای واقعی به مقدار مرجع نزدیک شود. در این مدل، دینامیک گرمایی سیستم به کمک ظرفیت حرارتی و تلفات حرارتی پایه مدل‌سازی شده است. نتایج در قالب سه نمودار شامل دمای واقعی و مرجع، توان ورودی اعمال‌شده و خطای ردیابی نمایش داده‌شده‌اند که به‌خوبی نحوه کنترل تطبیقی و پاسخ سیستم را در شرایط نویزی به تصویر می‌کشند.



شکل (۲): دمای کوره با کنترل مبتنی بر شبکه عصبی تطبیقی

یک سامانه کنترل تطبیقی مبتنی بر شبکه عصبی برای تنظیم دمای متغیر یک کوره قوس الکتریکی در شکل (۳) مدل‌سازی شده است.

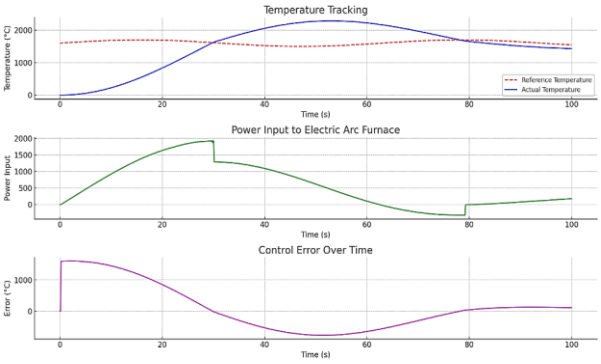


شکل (۳): دمای کوره با شبکه عصبی تطبیقی و مرجع متغیر زمانی

دمای مرجع به‌صورت سینوسی تغییر می‌کند تا شرایط عملیاتی واقعی‌تر شبیه‌سازی شود. کنترل‌کننده تطبیقی با استفاده از یک نورون مصنوعی با وزن‌های قابل به‌روزرسانی، سیگنال کنترل را تولید می‌کند و ورودی توان را برای پیروی از دمای مرجع تنظیم می‌نماید. مدل دینامیکی کوره با استفاده از ثابت زمانی و بهره سیستم مدل شده و به‌روزرسانی وزن‌ها نیز به روش مشابه گرادیان نزولی انجام می‌شود. نتایج شامل پیروی دقیق دمای واقعی از مقدار

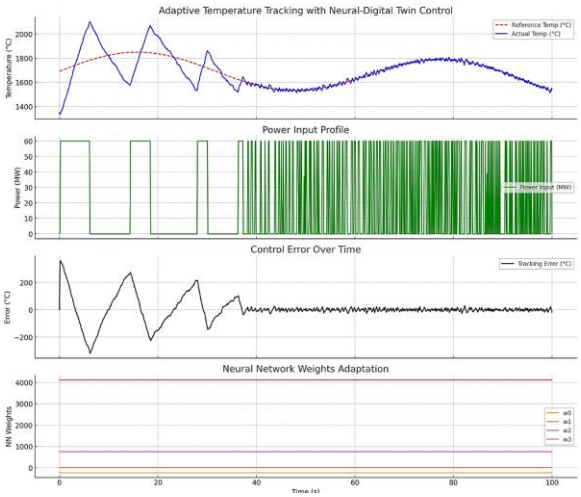
مرجع، تغییرات ورودی توان و میزان خطای کنترل، اثربخشی الگوریتم کنترل هوشمند را نشان می‌دهد.

شکل (٤) به مدل‌سازی یک کنترل‌کننده مبتنی بر شبکه عصبی مصنوعی برای پیگیری دمای یک کوره قوس الکتریکی می‌پردازد. دمای مرجع به‌صورت تابع سینوسی تغییر می‌کند تا شرایط پویا و پیچیده‌تری ایجاد گردد. شبکه عصبی شامل یک لایه پنهان با تابع فعال‌سازی سیگموئید و الگوریتم یادگیری گرادیان نزولی ساده برای به‌روزرسانی وزن‌هاست. ورودی شبکه شامل دمای فعلی، مشتق دما و خطای ردیابی است. در طول شبیه‌سازی، کنترل توان ورودی بر اساس خروجی شبکه انجام می‌شود و رفتار دینامیکی سیستم شبیه‌سازی‌شده و به صورت نمودارهایی از دمای واقعی، توان ورودی و خطای کنترلی نمایش داده‌شده است. این روش نشان می‌دهد که شبکه عصبی می‌تواند در شرایط دینامیکی متغیر، عملکرد کنترل را بهبود دهد.



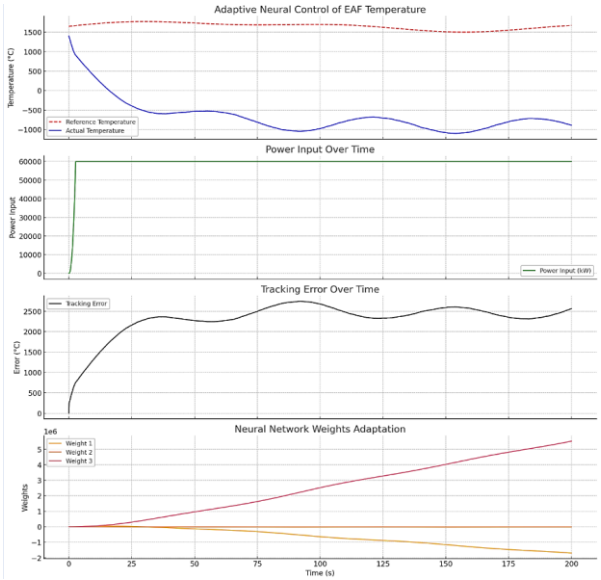
شکل (٤): کنترل تطبیقی دمای کوره با شبکه عصبی چندلایه

یک کنترل‌کننده تطبیقی مبتنی بر شبکه عصبی طراحی‌شده که با بهره‌گیری از بازخورد دوقلوی دیجیتال، توانایی بهینه‌سازی دمای کوره قوس الکتریکی را در برابر نوسانات شدید و افت انرژی حرارتی را در شکل (٥) نشان می‌دهد. سیستم کنترلی از چهار ورودی شامل دما، نرخ تغییر دما، خطا و انتگرال خطا بهره می‌گیرد و با یادگیری برخط وزن‌های شبکه عصبی، به‌صورت بلادرنگ با شرایط دینامیکی متغیر سازگار می‌شود. مدل دوقلوی دیجیتال نیز با در نظر گرفتن تلفات غیرخطی، نویز حرارتی و نوسانات خارجی، بازخوردی واقع‌گرایانه از وضعیت فیزیکی کوره فراهم می‌کند که منجر به دقت بالا در ردیابی دمای می‌شود.



شکل (٥): کنترل دما با یادگیری تطبیقی و دوقلوی دیجیتال

در شکل (٦)، کنترل تطبیقی دمای کوره قوس الکتریکی با بهره‌گیری از یک شبکه عصبی ساده و وزن‌های به‌روزشونده، برای ردیابی دقیق یک دمای مرجع پویا صورت گرفته است. تغییرات پیچیده در ورودی توان، تأثیر اتلاف انرژی نوسانی و وجود نویز محیطی، این سیستم را به محیطی واقعی نزدیک می‌کند. وزن‌های شبکه عصبی به صورت بلادرنگ و با الگوریتمی مشابه گرادیان نزولی به‌روزرسانی می‌شوند تا توانایی کنترل سیستم در برابر اختلالات افزایش یابد که نشان‌دهنده یک نمونه از پیاده‌سازی دوقلوی دیجیتال هوشمند در صنعت فولاد است.

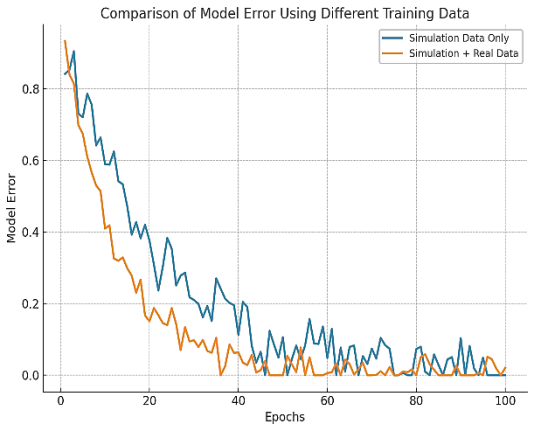


شکل (٦): کنترل با شبکه عصبی و دوقلوی دیجیتال

در مدل دیجیتال توسعه‌داده‌شده، خروجی اصلی بر پایه دما تعریف شده است. این دما نه تنها به‌عنوان شاخص کلیدی برای بازنمایی شرایط واقعی کوره قوس الکتریکی در فضای دیجیتال به‌کار می‌رود، بلکه نقش مرجع اصلی برای الگوریتم‌های کنترل تطبیقی و بهینه‌سازی انرژی ایفا می‌کند. به بیان دیگر، دوقلوی دیجیتال قادر است تغییرات لحظه‌ای دما را در شرایط عملیاتی مختلف شبیه‌سازی کرده و اطلاعات آن را به شبکه عصبی و الگوریتم‌های کنترلی منتقل نماید تا تصمیم‌گیری بهینه در خصوص مصرف انرژی و پایداری فرآیند ذوب صورت گیرد. همچنین برای جلوگیری از ساده‌سازی بیش‌ازحد، مدل علاوه بر دما، امکان استخراج پارامترهای جانبی مانند توان مصرفی و مشخصه‌های الکترود را نیز دارد، اما تمرکز اصلی در این پژوهش بر خروجی دما به‌عنوان متغیر حیاتی فرآیند قرارگرفته است.

برای اعتبارسنجی مدل پیشنهادی، از داده‌های شبیه‌سازی‌شده و واقعی استفاده شد. مجموعه داده واقعی از سیستم واحد مدیریت انرژی هوشمند [۲۷] به‌دست آمد، جایی که تغییرات دما و سیگنال‌های کنترلی در طول یک بازه ۳۰ روزه به‌طور پیوسته ثبت شدند. در راستای اعتبارسنجی مدل پیشنهادی، داده‌های واقعی به‌کار رفته در این پژوهش مشابه ساختار مجموعه‌های داده صنعتی معتبر است. به‌عنوان نمونه‌ای واقعی و مستند، در مجموعه داده ارائه‌شده توسط Engel و همکاران در مجله Scientific Data (سال ۲۰۲۵) [۲۷]، اطلاعات انرژی مصرفی و سیستم تهویه مطبوع یک محیط صنعتی هوشمند (به مدت چند سال) ثبت‌شده است. این مجموعه نمود وسیعی از رفتارهای واقعی کنترل، شرایط محیطی و مصرف انرژی را در سناریوهای متنوع پوشش می‌دهد. این داده‌ها شامل نویز اندازه‌گیری و نوسانات غیرمنتظره‌ای بودند که به‌طور طبیعی در محیط‌های واقعی رخ می‌دهند. سپس این مجموعه داده با پیش‌پردازش شامل نرمال‌سازی و حذف داده‌های پرت ترکیب شد و با داده‌های شبیه‌سازی تولیدشده در MATLAB/Simulink ادغام گردید. این مجموعه داده ترکیبی اجازه داد ابتدا شبکه عصبی با داده‌های شبیه‌سازی آموزش ببیند تا دینامیک سیستم را یاد بگیرد و سپس با استفاده از اختلالات واقعی تنظیم دقیق شود، که در نتیجه به افزایش پایداری و قابلیت اعتماد مدل منجر شد.

شکل (۷) روند تغییر خطای مدل را در دو سناریوی آموزشی مختلف نمایش می‌دهد. در حالت نخست که تنها داده‌های شبیه‌سازی‌شده برای آموزش استفاده شده‌اند، مشاهده می‌شود که کاهش خطا با شیب کمتر و همگرایی در مراحل بالاتری رخ می‌دهد. در مقابل، زمانی که مجموعه داده ترکیبی شامل داده‌های شبیه‌سازی و داده‌های واقعی به کار گرفته‌شده است، خطا در همان مراحل ابتدایی سریع‌تر کاهش‌یافته و در نهایت مقدار کمتری را نسبت به حالت صرفاً شبیه‌سازی تجربه می‌کند. این مقایسه نشان می‌دهد که افزودن داده واقعی به فرآیند آموزش نه‌تنها موجب بهبود دقت مدل می‌شود، بلکه سرعت همگرایی شبکه عصبی را نیز افزایش می‌دهد.



شکل (۷): مقایسه خطای مدل در آموزش صرفاً شبیه‌سازی‌شده و آموزش ترکیبی (شبیه‌سازی + واقعی)

جدول (۳) عملکرد روش پیشنهادی مبتنی بر شبکه عصبی و دوقلوی دیجیتال را در مقایسه با سه روش مرجع نشان می‌دهد. همان‌طور که مشاهده می‌شود، روش پیشنهادی در تمام معیارهای کلیدی مانند کاهش خطای ردیابی دما، مصرف انرژی کمتر، زمان نشست کوتاه‌تر و پایداری بالا در مواجهه با اغتشاشات، عملکرد بهتری از خود نشان داده است. همچنین این روش در مدیریت شرایط غیرخطی کوره‌های قوس الکتریکی نسبت به سایر روش‌ها از دقت و انعطاف‌پذیری بیشتری برخوردار است. این برتری‌ها نشان می‌دهند که استفاده همزمان از هوش مصنوعی و دوقلوی دیجیتال می‌تواند

بهینه‌سازی کنترل فرآیند را در محیط‌های صنعتی پیچیده تسهیل کند.

جدول (۳): مقایسه عددی عملکرد روش‌ها در کنترل دمای کوره قوس الکتریکی

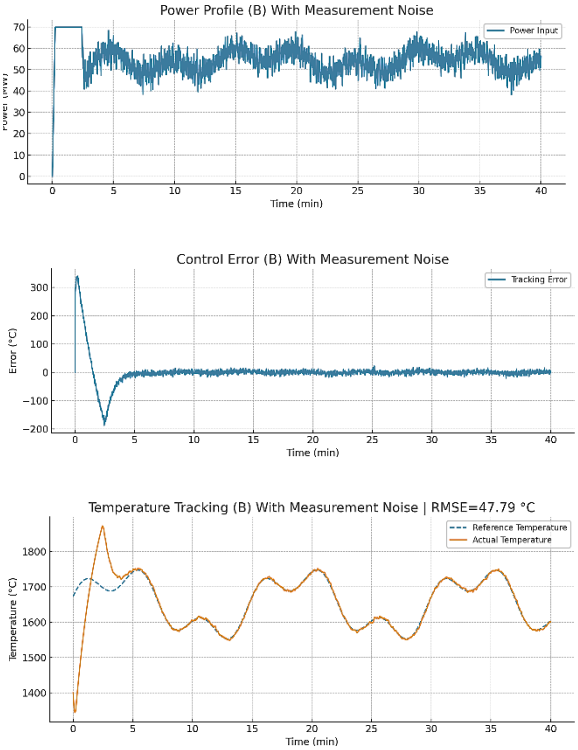
معیار ارزیابی	روش کلاسیک PID	کنترل تطبیقی PI	شبکه عصبی بدون دوقلو	روش پیشنهادی
میانگین خطای ردیابی دما	۷۴/۳	۴۸/۷	۳۵/۲	۱۲/۸
توان مصرفی کل	۵/۸۴	۵/۱۵	۴/۹۱	۴۳/۴
مقاومت به اغتشاشات	٪۷۲	٪۸۴	٪۸۸	٪۹۵
زمان نشست سیستم	۶۱	۴۵	۳۳	۲۱
نرخ سازگاری با بارهای غیرخطی	٪۶۸	٪۸۱	٪۸۷	٪۹۷

با توجه به جدول (۴)، مدل پیشنهادی در تمامی شاخص‌ها نسبت به روش‌های مرجع عملکرد بهتری ارائه کرده است. مقدار MSE و RMSE کمتر بیانگر دقت بالاتر مدل در پیش‌بینی مقادیر واقعی است. همچنین کاهش MAE نشان‌دهنده کاهش انحراف مطلق خطا در تخمین‌ها است. از سوی دیگر، افزایش ضریب تعیین (R^2) تا مقدار ۱۸۹/. نشان می‌دهد که مدل پیشنهادی قادر است بیش از ۹۸ درصد تغییرات داده‌های واقعی را توضیح دهد، که نسبت به سایر روش‌ها بهبود قابل‌توجهی محسوب می‌شود.

جدول (۴): جدول نمونه ارزیابی عددی

روش مقایسه	MSE[1]	RMSE	MAE[3]	R^2
روش RLS	۰/۰۱۵	۰/۱۲۲	۰/۰۸۹	۰/۹۴۱
روش EKF	۰/۰۱۱	۰/۱۰۵	۰/۰۷۳	۰/۹۵۶
روش NN-Estimator	۰/۰۰۹	۰/۰۹۷	۰/۰۶۸	۰/۹۶۳
مدل پیشنهادی	۰/۰۰۵	۰/۰۷۱	۰/۰۵۱	۰/۹۸۱

در راستای اعتبارسنجی دقیق‌تر روش پیشنهادی، چند سناریوی متنوع و نزدیک به شرایط واقعی عملیاتی مورد بررسی قرار گرفت. نخست در شکل (۸)، اثر نویز اندازه‌گیری بر سیگنال‌های دما و انرژی ورودی لحاظ شد. در چنین شرایطی، الگوریتم باید توانایی جداسازی مؤلفه‌های مفید از اغتشاشات تصادفی را داشته باشد تا عملکرد کنترل و پیش‌بینی مختل نشود. نتایج نشان داد که حتی با افزایش سطح نویز، تخمین مقادیر اصلی دما و انرژی همچنان در محدوده قابل قبول باقی می‌ماند.



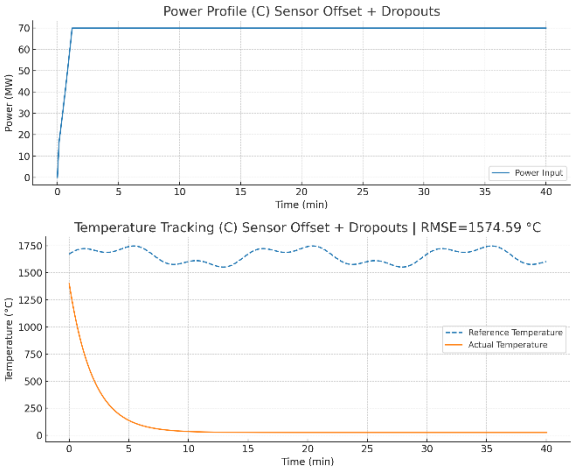
شکل (۸): بررسی وجود نویز اندازه‌گیری در سیگنال‌های دما و انرژی ورودی

در سناریوی دوم، خطاهای حسگر به صورت آفست ثابت و همچنین قطع متناوب داده‌ها درشکل (۹) شبیه‌سازی گردید. بروز این دسته خطاها در محیط‌های واقعی امری اجتناب‌ناپذیر است و می‌تواند باعث ایجاد انحراف در تصمیم‌گیری‌های کنترلی شود. روش ارائه‌شده توانست با استفاده از سازوکارهای

[1] Mean Squared Error

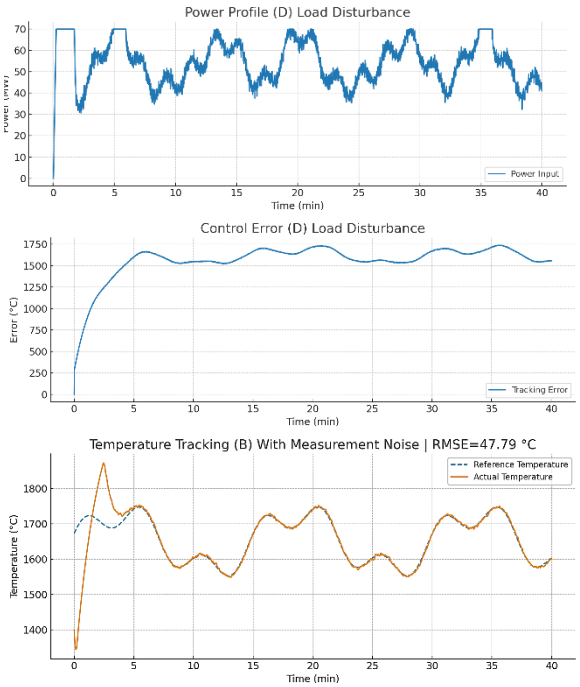
[2] Root Mean Squared Error
[3] Mean Absolute Error

جبرانی تعبیه‌شده، اثر این خطاها را کاهش داده و پایداری سیستم را حفظ نماید.



شکل (۹): بررسی بروز خطاهای حسگر (آفست و قطع مقطعی داده‌ها)

در ادامه، تغییرات شرایط عملیاتی نظیر نوسان در بار حرارتی و تغییر نرخ تزریق انرژی در شکل (۱۰) بررسی شد.



شکل (۱۰): تغییرات شرایط عملیاتی شامل نوسان در بار و تغییر نرخ تزریق انرژی

این تغییرات اغلب موجب دینامیک‌های سریع و غیرخطی در سیستم می‌شوند که برای بسیاری از روش‌های متعارف چالش‌برانگیز است. با این حال، نتایج شبیه‌سازی نشان داد که رویکرد پیشنهادی توانسته است بدون بروز ناپایداری، مقادیر دما و انرژی را در سطح مطلوب کنترل نماید و انحرافات گذرا را به سرعت میرا کند.

ترکیب این آزمایش‌ها به خوبی نشان داد که چارچوب طراحی‌شده نه تنها در شرایط ایده‌آل بلکه در حضور اغتشاشات محیطی، خطاهای سنسوری و تغییرات بار نیز کارآمد است. بدین ترتیب، روش پیشنهادی از مقاومت بالا و قابلیت اطمینان برخوردار بوده و می‌تواند به عنوان یک راهکار عملی برای کاربردهای واقعی مورد استفاده قرار گیرد. به منظور بررسی میزان تأثیرپذیری عملکرد چارچوب پیشنهادی از تغییرات پارامترهای کلیدی، تحلیل حساسیت در سه محور اصلی شامل نرخ یادگیری شبکه عصبی، پارامترهای کنترلی سیستم و داده‌های ورودی دوقلوی دیجیتال طبق جدول (۵) صورت گرفت.

۱- نرخ یادگیری شبکه عصبی:

آزمایش‌ها نشان دادند که مقادیر پایین نرخ یادگیری سبب کندی همگرایی و افزایش زمان آموزش می‌شود، در حالی که مقادیر بیش‌ازحد بزرگ منجر به نوسان و کاهش دقت در فرآیند تطبیق مدل می‌گردد. محدوده بهینه‌ای از نرخ یادگیری شناسایی شد که در آن شبکه ضمن حفظ سرعت یادگیری، از پایداری بالایی برخوردار است.

۲- پارامترهای کنترلی:

تغییر ضرایب کنترلی نظیر بهره‌های تناسبی و انتگرالی بر پایداری و دقت ردیابی مرجع اثرگذار است. نتایج نشان دادند که سیستم در برابر تغییرات کوچک این پارامترها مقاوم بوده و پایداری خود را حفظ می‌کند، اما در صورت افزایش بیش‌ازحد این ضرایب، پاسخ سیستم دچار اُورشوت بالا و نوسانات گذرا می‌شود. در مقابل، کاهش شدید آن‌ها باعث کند شدن پاسخ دینامیکی خواهد شد.

۳- داده‌های ورودی دوقلوی دیجیتال:

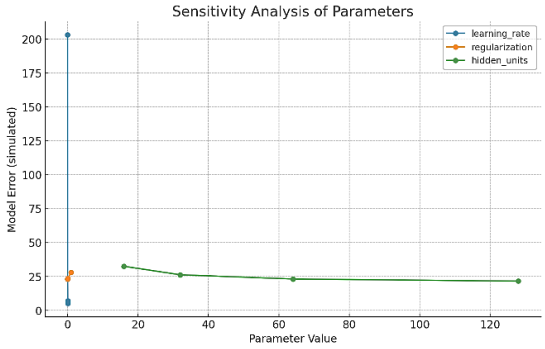
کیفیت و دقت داده‌های ورودی نقش مهمی در کارایی کل سامانه دارند. آزمایش‌ها بیانگر آن بود که مدل پیشنهادی قادر است در برابر نویزهای کم تا متوسط عملکرد مناسب خود را حفظ کند. با این حال، در شرایطی که داده‌های ورودی دارای خطاهای قابل‌توجه

یا تأخیر زیاد باشند، کاهش محسوس در دقت تخمین و پیش‌بینی مشاهده شد. برای مقابله با این موضوع، از مکانیزم‌های فیلترینگ و هموارسازی داده‌ها استفاده گردید.

جدول (۵): مقایسه تحلیل حساسیت

پارامتر مورد تغییر	محدوده تغییر (٪)	شاخص عملکرد اصلی	تغییر در پایداری سیستم
نرخ یادگیری شبکه عصبی	۰/۰۰۱ تا ۰/۵۰۰	۰/۰۱۲← ۰/۰۲۸	پایدار، افت جزئی در یادگیری سریع
پارامتر کنترلی K_p	٪۳۰- تا ٪۳۰	۰/۰۱۰ → ۰/۰۱۸	پایدار در بیشتر محدوده‌ها
پارامتر کنترلی K_i	٪۳۰- تا ٪۳۰	۰/۰۰۹ → ۰/۰۲۱	پایدار، اما کندی در پاسخ
داده ورودی دوقلوی دیجیتال	٪۰ تا ٪۱۰ نویز افزوده	۰/۰۱۱ → ۰/۰۳۲	پایدار، افزایش خطای تخمین

نمودار حساسیت ارائه‌شده در شکل (۱۱) میزان تأثیرگذاری تغییرات هر یک از پارامترهای کلیدی مدل بر خطای کلی سیستم را نشان می‌دهد. همان‌طور که مشاهده می‌شود، برخی از پارامترها در مقایسه با سایرین حساسیت بالاتری داشته و تغییر اندک در مقدار آن‌ها منجر به افزایش قابل‌توجه خطای مدل می‌شود.



شکل (۱۱): تحلیل حساسیت پارامترها بر خطای مدل

این تحلیل می‌تواند در فرآیند بهینه‌سازی و انتخاب استراتژی‌های کنترلی به‌عنوان راهنمایی مهم مورد استفاده قرار گیرد، چرا که تمرکز بر پارامترهای حساس‌تر موجب بهبود کارایی سیستم و افزایش دقت تخمین خواهد شد.

۶- نتیجه‌گیری

در این مقاله، یک چارچوب نوآورانه برای کنترل هوشمند و بهینه‌سازی مصرف انرژی در کوره‌های قوس الکتریکی در صنعت فولاد ارائه شد. این چارچوب مبتنی بر ترکیب شبکه عصبی تطبیقی و دوقلوی دیجیتال است که با هدف افزایش دقت ردیابی دما، کاهش مصرف انرژی و بهبود پایداری سیستم طراحی‌شده است. نتایج شبیه‌سازی‌ها نشان داد که الگوریتم پیشنهادی نه‌تنها توانسته است با دقت بالا رفتار حرارتی کوره را دنبال کند، بلکه با سازگاری با شرایط متغیر، اغتشاشات محیطی و مدل‌سازی ضرایب اتلاف حرارتی، نسبت به سایر روش‌های کنترل کلاسیک و تطبیقی عملکرد بهتری از خود نشان داده است. استفاده از دوقلوی دیجیتال باعث شد تا بازخورد دقیق‌تری از وضعیت حرارتی واقعی سیستم به‌دست آید و این بازخورد نقش مهمی در آموزش برخط شبکه عصبی و اصلاح وزن‌های آن ایفا کرد. همچنین، ساختار تطبیقی کنترل باعث شد تا سیستم نسبت به تغییرات تدریجی و ناگهانی در دینامیک کوره مقاوم باشد. در مسیر توسعه این تحقیق، به‌جای محدود کردن بحث به شبکه‌های کانولوشنی، می‌توان از ساختارهای پیچیده‌تر یادگیری عمیق مانند شبکه‌های بازگشتی پیشرفته یا مدل‌های هیبریدی ترکیبی بهره گرفت تا هم وابستگی‌های زمانی و هم در صورت وجود بعد مکانی در داده‌ها به شکل دقیق‌تری مدل‌سازی شوند.

همچنین پیاده‌سازی الگوریتم پیشنهادی در محیط واقعی کارخانه و مقایسه عملکرد آن با داده‌های عملیاتی واقعی می‌تواند گامی مؤثر در تجاری‌سازی این رویکرد باشد. از دیگر مسیرهای توسعه، می‌توان به استفاده از یادگیری تقویتی برای تنظیم خودکار سیاست کنترلی و گسترش چارچوب به سیستم‌های چندورودی-چندخروجی اشاره کرد. همچنین، یکپارچه‌سازی این سیستم با زیرساخت‌های اینترنت اشیا و تحلیل کلان‌داده‌های صنعتی برای پیش‌بینی نگهداری و کاهش زمان توقف نیز قابل بررسی است.

References

[1] B. Yu, Y. Dai, J. Fu, J. Qi, and X. Li, "Industrial risks assessment for the large-scale development of electric arc furnace steelmaking technology," J Applied Energy, vol. 377, p. 124611, 2025.

[2] L. Holappa, "A general vision for reduction of energy consumption and CO2 emissions from the steel industry," J Metals, vol. 10, no. 9, p. 1117, 2020.

[3] E. Carpanzano and D. Knüttel, "Advances in artificial intelligence methods applications in industrial control systems: Towards cognitive self-optimizing manufacturing systems," J Applied sciences, vol. 12, no. 21, p. 10962, 2022.

[4] H. Jiang, S. Qin, J. Fu, J. Zhang, and G. Ding, "How to model and implement connections between physical and virtual models for digital twin application," J Journal of Manufacturing Systems, vol. 58, pp. 36-51, 2021.

[5] L. Cui, Y. Xiao, D. Liu, and H. Han, "Digital twin-driven graph domain adaptation neural network for remaining useful life prediction of rolling bearing," J Reliability Engineering System Safety, vol. 245, p .109991, 2024.

[6] M. M. Abadi, H. Tang, and M. M. Rashidi, "A review of simulation and numerical modeling of electric arc furnace (EAF) and its processes," J Heliyon, vol. 10, no. 11, 2024.

[7] B. Tian, G. Wei, H. Hu, R. Zhu, H. Bai, Z. Wang, and L. Yang, "Effects of fuel injection and energy efficiency on the production and environmental parameters of electric arc furnace-heat recovery systems," J Journal of Cleaner Production, vol. 405, p. 136909, 2023.

[8] S. Jawahery, V.-V. Visuri, S. O. Wasbø, A. Hammervold, N. Hyttinen, and M. Schlautmann, "Thermophysical model for online optimization and control of the electric arc furnace," J Metals, vol. 11, no. 10, p. 1587, 2021.

[9] L. Shiqi, "Electric Arc Furnace for Steelmaking," in The ECPH Encyclopedia of Mining and Metallurgy: Springer, 2024, pp. 548-550.

[10] I. A. Abbas and M. K. Mustafa, "A review of adaptive tuning of PID-controller: Optimization techniques and applications," J International Journal of Nonlinear Analysis Applications, vol. 15, no. 2 ,pp. 29-37, 2024.

[11] N. Ajlouni, A. Osyavas, F. Ajlouni, A. Almassri, F. Takaoglu, and M. Takaoglu, "Enhancing PID control robustness in CSTRs: a hybrid approach to tuning under external disturbances with GA, PSO, and machine learning," J Neural Computing Applications, vol. 37, no. 18, pp. 12153-12177, 2025.

[12] Kiheha, M. M., & Barahouei, S. (2024). A novel approach to graph dimensionality reduction based on deep learning using fuzzy logic and random walks. Theoretical and Applied Machine Intelligence Research, 2(1), 131–141.

[13] R. G. Alarcón, M. A. Alarcón, A. H. González, and A. Ferramosca, "Artificial Neural Networks for Energy Demand Prediction in an Economic MPC-Based Energy Management System," J International Journal of Robust Nonlinear Control, vol. 35, no. 2, pp. 642 - 658 ,2025.

[14] Y.-P. Chen, V. Karkaria, Y.-K. Tsai, F. Rolark, D. Quispe, R. X. Gao, J. Cao, and W. Chen, "Real-time decision-making for Digital Twin in additive manufacturing with Model Predictive Control using time-series deep neural networks," J Journal of Manufacturing Systems, vol. 80, pp. 412-424, 2025.

[15] A. Sivtsov, O. Y. Sheshukov, D. Egiazar'yan, M. Tsymbalist, and P. Orlov, "Automated process control in electric arc furnaces in the aspect of digital twin technology," J Izvestiya. Ferrous Metallurgy, vol. 67, no. 4, pp. 481-489, 2024.

[16] M. M. Abadi, H. Tang, and M. M. Rashidi, "A review of simulation and numerical modeling of electric arc furnace (EAF) and its processes," J Heliyon, 2024.

[17] Dehghani Mahmoudabadi, M., Mirzaei, K., & Peirovi, F. (2023). Improving credit assignment of rules in learning classifier systems using Markov reinforcement learning for secondary protein structure prediction. Theoretical and Applied Machine Intelligence Research, 1(2), 92–104.

[18] Y. Yin, Z. Wang, L. Zheng, Q. Su, and Y. Guo, "Autonomous UAV navigation with adaptive control based on deep reinforcement learning," J Electronics, vol. 13, no. 13, p. 2432, 2024.

[19] A. A. Khater, M. Fekry, M. El-Bardini, and A. M. El-Nagar, "Deep reinforcement learning-based adaptive fuzzy control for electro-hydraulic servo system," J Neural Computing Applications, pp. 1-18, 2025.

[20] A. Barzegar and D.-J. Lee, "Deep reinforcement learning-based adaptive controller for trajectory tracking and altitude control of an aerial robot," J Applied Sciences, vol. 12 ,no. 9, p. 4764, 2022.

[21] B. Gajdzik, Digital Transformation Towards Smart Steel Manufacturing. Springer, 2025.

[22] B. Zheng, M. Pan, Q. Liu, X. Xu, C. Liu, X. Wang, W. Chu, S. Tian, J. Yuan, and Y. Xu, "Data-driven assisted real-time optimal control

strategy of submerged arc furnace via intelligent energy terminals considering large-scale renewable energy utilization," J Scientific Reports, vol. 14, no. 1, p. 5582, 2024.

[23] M. Jafari, A. Kavousi-Fard, T. Chen, and M. Karimi, "A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future," J IEEe Access, vol. 11, pp. 17471-17484, 2023.

[24] H. Hosamo, M. H. Hosamo, H. K. Nielsen, P. R. Svennevig, and K. Svidt, "Digital Twin of HVAC system (HVACDT) for multiobjective optimization of energy consumption and thermal comfort based on BIM framework with ANN-MOGA," J Advances in building energy research, vol. 17, no. 2, pp. 125-171, 2023.

[25] J. D. Hernández, L. Onofri, and S. Engell, "Modeling and energy efficiency analysis of the steelmaking process in an electric arc furnace," J Metallurgical Materials Transactions B, vol. 53, no. 6, pp. 3413-3441, 2022.

[26] T. R. Mohan, R. A. Uthra, and J. P. Roselyn, "Intelligent and effective means of power utilization in induction furnace controlled melting process of foundry industries reducing CO2 emissions," J Electrical Engineering, pp. 1-20, 2025.

[27] J. Engel, A. Castellani, P. Wollstadt, F. Lanfermann, T. Schmitt, S. Schmitt, L. Fischer, S. Limmer, D. Luttropp ,and F. Jomrich, "A real-world energy management dataset from a smart company building for optimization and machine learning," J Scientific Data, vol. 12, no. 1, p. 864, 2025.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 53 - 71

Smart Energy Optimization and Adaptive Control of Electric Arc Furnaces in the Steel Industry using Digital Twin and Neural Networks

Sara Mahmoudi Rashid*

Instructor of Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

Article Information

Original Research Paper

Received:
2025 May 16

Accepted:
2025 September 30

Keywords:
Industrial Artificial Intelligence, Electric Arc Furnace, Energy Consumption Optimization, Adaptive Control, Learning Neural Network, Steel Industry Efficiency

Corresponding Author*:
s.mahmoudirashid@tabrizu.ac.ir

Abstract

This study proposes a novel framework for energy optimization and adaptive control of Electric Arc Furnaces (EAFs) in the steel industry, based on the integration of a digital twin model and adaptive neural networks. The main objective is to reduce energy consumption, enhance process stability, and improve control accuracy under varying operational conditions. To achieve this, a real-time updated digital twin of the steel melting process was first developed. Subsequently, adaptive neural networks were employed to estimate the nonlinear dynamics of the system and design an intelligent controller. The proposed system was implemented in the MATLAB/Simulink environment and evaluated under multiple operational scenarios. Simulation results demonstrate that the proposed approach achieves an average energy consumption reduction of 8.7%, a 24.3% improvement in transient response, and a 14.12% increase in power tracking accuracy compared to conventional controllers. These outcomes highlight the effectiveness of the proposed method in enhancing energy efficiency and operational stability in thermal processes within the steel industry and suggest its potential as a scalable model for smart system implementation in other industrial applications.

: 10.22034/ABMIR.2025.23151.1128

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۸۸ - ۷۳

طبقه‌بندی عیوب سطح فولاد با مدلی بهبودیافته از معماری VGG16

طوبی ترابی پور [۱]، ابوالفضل گندمی [۲*]

[۱] دانشجوی دکتری مهندسی کامپیوتر، دانشگاه آزاد اسلامی، یزد، ایران

[۲] استادیار دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، یزد، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۶/۱۴

تاریخ پذیرش:
۱۴۰۴/۰۸/۱۶

کلیدواژه‌ها:
یادگیری عمیق، طبقه‌بندی عیوب سطح فولاد، VGG16، مکانیسم توجه

نویسنده‌مسئول:
Abolfazl.Gandomi@iau.ac.ir

چکیده

ورق‌های فولادی به‌عنوان یکی از اجزای کلیدی در صنایع خودروسازی مورد استفاده قرار می‌گیرند. طبقه‌بندی دقیق عیوب سطحی این ورق‌ها، نقشی اساسی در تضمین کیفیت آن‌ها دارد. چالش اصلی در این زمینه، شیوه بازرسی ورق‌های فولادی است زیرا، روش‌های سنتی بازرسی با محدودیت‌های قابل‌توجه ناشی از خطای انسانی مواجه‌اند. هرچند در سال‌های اخیر، الگوریتم‌های یادگیری عمیق به‌عنوان رویکردی مؤثر و خودکار به‌منظور شناسایی عیوب مطرح شده‌اند اما طبقه‌بندی عیوب کوچک و پیچیده، همچنان دشوار باقی‌مانده و نیازمند مدل‌سازی بهینه اطلاعات مکانی و کانالی است. در این پژوهش، مدلی نوین مبتنی بر معماری VGG16 و یادگیری انتقالی ارائه می‌شود که با بهره‌گیری از ماژول توجه فضایی دوبُعدی و مکانیزم بزرگ‌نمایی مبتنی بر توجه بر کاهش نویز پس‌زمینه و تمرکز بر نواحی کلیدی تصویر تمرکز دارد. پس از ارزیابی مدل بر روی دو مجموعه داده NEU-CLS و NEU-DET به‌ترتیب دقت ۹۹/۹۸ و ۹۹/۹۲ درصد به‌دست آمد که نسبت به مدل پایه VGG16 (حداقل ٤ درصد) و نسبت به پژوهش‌های پیشین، بهبود (حداقل ۰/۰۱۸ درصد) داشته است. این نتایج، کارایی بالای مدل را در طبقه‌بندی دقیق عیوب سطح فولاد نشان می‌دهد. در پایان، جهت بهره‌برداری عملی، یک وب اپلیکیشن توسعه‌یافته و در اختیار متخصصان قرارگرفته است.

: 10.22034/ABMIR.2025.23563.1162

E-ISSN: 2821-2037

۱- مقدمه

صنعت فولاد به‌عنوان یکی از پایه‌های اصلی اقتصاد در ایران شناخته می‌شود و محصولات متنوع آن در بخش‌های مختلفی همچون ساخت‌وساز و خودروسازی کاربرد گسترده‌ای دارند. در این میان، ورق‌های فولادی جایگاه ویژه‌ای داشته و کیفیت سطح آن‌ها به‌طور مستقیم بر کارایی و ایمنی محصول نهایی تأثیرگذار است. عیوب سطحی ورق‌های فولادی، می‌توانند در اثر عواملی نظیر ترکیب شیمیایی نامناسب و یا مشکلات فرایند تولید، ایجاد شوند و درصورت عدم شناسایی به‌موقع، موجب کاهش بهره‌وری و افت کیفیت محصول خواهند شد. ازاین‌رو، توسعه روش‌هایی به‌منظور طبقه‌بندی سریع و دقیق این عیوب از اهمیت بالایی برخوردار است. به‌منظور بررسی و شناسایی عیوب سطحی این ورق‌ها، دو رویکرد اصلی وجود دارد:

روش‌های سنتی بازرسی: این روش‌ها، معمولاً شامل بازرسی چشمی توسط سرکارگران بخش تولید هستند. هرچند ساده و متداول‌اند اما با محدودیت‌های قابل‌توجهی مانند زمان‌بر بودن، نیاز به نیروی انسانی زیاد و احتمال بروز خطای انسانی همراه‌اند. چنین محدودیت‌هایی موجب می‌شود؛ این روش‌ها توانایی پاسخ‌گویی به نیازهای تولید انبوه ورق‌های فولادی را نداشته باشند [۱].

روش‌های مبتنی بر یادگیری عمیق: در سال‌های اخیر، الگوریتم‌های یادگیری عمیق به‌عنوان جایگزینی نوین و کارآمد برای روش‌های سنتی معرفی‌شده‌اند. این الگوریتم‌ها، با قابلیت یادگیری و استخراج خودکار ویژگی‌ها، قادرند عیوب سطحی ناشی از اشکال و اندازه‌های متنوع را با دقت بالا و بدون تأثیرپذیری از عوامل محیطی شناسایی کنند. ازجمله مدل‌های موفق در این زمینه می‌توان به شبکه‌های عصبی کانولوشنی (CNN[1])، الگوریتم‌های تشخیص اشیا مانند Faster R-CNN و خانواده الگوریتم‌های YOLO[2] اشاره کرد [۲].

در سال‌های اخیر، برای افزایش دقت این مدل‌ها، استفاده از مکانیسم‌های توجه به‌صورت مرسوم در عرصه بین‌المللی جا افتاده است. افزودن مکانیسم توجه به مدل، این قابلیت را میسر می‌سازد که هوشمندانه بر روی نواحی مهم‌تر در تصویر متمرکز شده و ویژگی‌های متمایزکننده‌تر را تقویت کند. این امر، منجر به حذف نویز پس‌زمینه می‌شود [۳].

علی‌رغم این پیشرفت‌ها، یک خلأ تحقیقاتی به‌شکل بارز خودنمایی می‌کند. تمرکز مکانیسم‌های توجه موجود، بر یک بُعد از اطلاعات (مانند توجه کانالی در SE[3] و عدم امکان مدل‌سازی بهینه با اطلاعات مکانی و ویژگی‌های کانالی است. این محدودیت‌ها، می‌تواند منجر به کاهش دقت شود. ازسوی‌دیگر، ایجاد ترکیب بهینه ویژگی‌ها از سطوح مختلف که برای طبقه‌بندی هم‌زمان عیوب بزرگ و کوچک ضروری است، یک چالش جدی محسوب می‌شود. برای رفع این چالش‌ها، در این تحقیق تلاش می‌شود؛ یک معماری نوین مبتنی بر توجه فضایی برای طبقه‌بندی دقیق و بهینه عیوب سطح ورق‌های فولاد ارائه شود. در این تحقیق، یک مدل سه‌مرحله‌ای شامل استخراج ویژگی‌ها، ماژول توجه و طبقه‌بندی ارائه می‌شود:

مرحله اول؛ استخراج ویژگی با یادگیری انتقال و معماری VGG16[4].

مرحله دوم؛ ایجاد یک نگاشت توجه فضایی دوبُعدی و یک فرایند (بزرگ‌نمایی مجدد مبتنی بر توجه).

مرحله سوم؛ طبقه‌بندی نوع خرابی در ورق‌های فولاد و نمایش آن در وب اپلیکیشن ایجاد شده.

در بخش ۲ این مقاله، به کارهای مرتبط، در بخش ۳ به روش تحقیق و در بخش ٤ به شبیه‌سازی و ارزیابی پرداخته می‌شود. در بخش ۵ به ساخت وب اپلیکیشن کاربردی، در بخش ٦ به مقایسه مدل‌های طبقه‌بندی عیوب سطح فولاد و در بخش ۷ به نتیجه‌گیری و کارهای آینده پرداخته می‌شود.

۲- کارهای مرتبط

از سال ۲۰۱۵ میلادی تاکنون، تحقیقات بسیاری انجام شده است. رن و همکارانش [٤]، در سال ۲۰۱۵ با تمرکز بر طبقه‌بندی اشیا در زمان واقعی، معماری Faster R-CNN را معرفی کردند. آن‌ها

[1] Conveloutional neural network
[2] You Only Look Once
[3] Squeeze-And-Excitation
[4] Visual Geometry Group

جهت ارزیابی از مجموعه داده PASCAL VOC 2007 استفاده کردند و به MAP[1] (میانگین دقت متوسط) برابر با ۷۳/۲ درصد دست یافتند. هی و همکارانش [۵]، در سال ۲۰۱۹ یک رویکرد انتها–به–انتها به‌منظور طبقه‌بندی عیوب سطح فولاد ارائه دادند که در آن، ویژگی‌های سلسله‌مراتبی از سطوح مختلف شبکه با هم ترکیب می‌شدند. آن‌ها جهت ارزیابی از مجموعه داده NEU-DET[2] که توسط خودشان معرفی شد، استفاده کردند و در بخش طبقه‌بندی عیوب به‌دقت برابر با ۹۹/۷ درصد رسیدند. ژانگ و یو [۶]، در سال ۲۰۲۰ مدل RetinaNet را با افزودن یک ماژول توجه کانالی تفاضلی و یک ماژول ترکیب ویژگی فضایی تطبیقی (ASFF[3]) بهبود دادند. آن‌ها این مدل را بر روی مجموعه داده NEU-DET ارزیابی کرده و به MAP برابر با ۷۸/۲۵ درصد دست یافتند.

وی و همکارانش [۷] در سال ۲۰۲۰ معماری Faster R-CNN را با استفاده از یک لایه استخر ناحیه موردعلاقه وزن‌دار و کانولوشن‌های تغییرپذیر برای تشخیص بهتر عیوب نامنظم بهینه کردند. در ارزیابی انجام شده بر روی مجموعه داده NEU-DET، آن‌ها به MAP برابر با ۷۲/۴ درصد رسیدند. تانگ و همکارانش [۸] در سال ۲۰۲۱ از یک مدل مبتنی بر ResNet50 به‌همراه مکانیسم توجه و یک ماژول ادغام چندمقیاسی (MSMP[4]) استفاده کردند. این روش که بر روی مجموعه داده NEU-DET ارزیابی شد، توانست بهبود MAP به میزان ۳/۶۵ درصد نسبت به شبکه پایه را نشان دهد. گو و همکارانش [۹] در سال ۲۰۲۲ مدل MSFT-YOLO را معرفی کردند که نسخه‌ای بهبودیافته از YOLOv5 با یک جزء مبتنی بر ترنسفورمر به‌منظور ترکیب اطلاعات سراسری و محلی است. این مدل بر روی مجموعه داده NEU-DET ارزیابی شد و به میانگین دقت MAP برابر با ۷۵/۲ درصد دست‌یافت. لی و همکارانش [۱۰] در سال ۲۰۲۴ یک مدل مبتنی بر YOLOv5s با ماژول‌های نوآورانه MSFE[5] (استخراج ویژگی چندمقیاسی) و EFF[6] (ترکیب کارآمد ویژگی) ارائه کردند. این مدل بر روی مجموعه داده NEU-DET با معیار MAP ارزیابی شد و به دقت ۷۳/۰۸ درصد رسید. گائو و همکارانش [۱۱] در سال ۲۰۲۴ مدل YOLOv5-KBS را با ترکیب الگوریتم YOLOv5، مکانیسم توجه SE و شبکه ترکیب ویژگی BiFPN[7] ارائه دادند. در ارزیابی انجام شده بر روی مجموعه داده NEU-DET، این مدل به MAP برابر با ۷۹/۹ درصد و سرعت ۷۰ فریم بر ثانیه دست‌یافت.

ژنگ و همکارانش [۱۲] در سال ۲۰۲۵، مدلی مبتنی بر ترکیب تبدیل چندموجک و شبکه خودرمزگذار ارائه دادند. این مدل با استخراج ویژگی‌های فرکانسی و مکانی توسط چندموجک و کاهش ابعاد ویژگی‌ها با خودرمزگذار، به طبقه‌بندی دقیق عیوب سطح فولاد دست‌یافته است. در ارزیابی انجام شده بر روی مجموعه داده NEU-CLS[8]، این مدل به دقت برابر با ۹۹/۴۴ درصد، با SVM و ۹۹/۰۶ درصد با BPNN دست‌یافت.

وانگ و همکارانش [۱۳] در سال ۲۰۲۵ یک مدل مبتنی بر YOLOv11n با ماژول‌های نوآورانه AMSPPF[9] (ادغام تطبیقی چندمقیاسی سریع)، C2PSA-DSAM[10] (ماژول توجه فضایی تغییرپذیر)، LDConv[11] (کانولوشن خطی تغییرپذیر) و تابع هزینه Inner-CIoU[12] ارائه دادند. در ارزیابی انجام شده بر روی مجموعه داده NEU-DET، این مدل به MAP برابر با ۸۳/۰ درصد دست‌یافت. لی و همکاران [۱۴] در سال ۲۰۲۵ یک مدل مبتنی بر الگوریتم چندمقیاسی و تقویت ویژگی چندبُعدی ارائه دادند. این مدل از ماژول MRAM[13] (تجمیع باقی‌مانده چندشاخه‌ای) برای استخراج هم‌زمان ویژگی‌ها و از ماژول MAEH[14] (سر تقویت توجه چندبُعدی) برای ادغام اطلاعات در وجوه مختلف مقیاس و کانال استفاده می‌کند. این مدل بر روی مجموعه داده NEU-DET

[1] Mean Average Precision
[2] Northeastern University Surface Defect Detection
[3] Adaptive Spatial Feature Fusion
[4] Multi-Scale Merging/Pooling Module
[5] Multi-Scale Feature Extraction
[6] Efficient Feature Fusion
[7] Bidirectional Feature Pyramid Network
[8] Northeastern University-Classification
[9] Adaptive Multi-Scale Pooling And Fusion
[10] Deformable Spatial Attention Module With Cross-Stage Partial Spatial Attention
[11] Linear Deformable Convolution
[12] Inner Complete Iou
[13] Multi-Branch Residual Aggregation Module
[14] Multi-Dimensional Attention Enhancement Head

با معیار MAP ارزیابی شد و به دقت برابر با ۸۲/۷ درصد و سرعت ۸۹ FPS رسید. ژنگ و همکاران [۱۵] در همان سال، یک مدل شبکه عصبی کانولوشنی سبک مبتنی بر ترکیب تبدیل چندموجکی (LWT[1]) و مکانیزم توجه ویژگی‌ها (FAG[2]) ارائه کردند. این مدل بر روی مجموعه داده NEU-CLS با معیار دقت ارزیابی شد و به دقت برابر با ۹۹/۸۰ درصد رسید.

۳- روش تحقیق

در این پژوهش، یک مدل نوین مبتنی بر معماری قدرتمند VGG16 و شبکه عصبی کانولوشنی (CNN) برای طبقه‌بندی دقیق عیوب سطح فولاد ارائه می‌شود. این مدل از رویکرد یادگیری انتقالی در معماری VGG16 بهره برده و با فریز کردن لایه انتهایی این شبکه، از آن به‌عنوان استخراج کننده ویژگی‌های اصلی استفاده می‌کند. اما در بخش اصلی مدل از یک شبکه عصبی کانولوشن تغییریافته با یک ماژول توجه فضایی و یک مکانیزم بزرگ‌نمایی مجدد برای افزایش دقت و پایداری مدل که به‌اختصار «ماژول توجه» نامیده شده، استفاده‌شده است. در بخش پایانی نیز طبقه‌بندی نهایی نوع آسیب ورق‌های فولاد انجام می‌شود. معماری کلی مدل شامل سه بخش اصلی است: استخراج ویژگی، ماژول توجه و طبقه‌بندی نهایی.

۳-۱ معماری مدل

۱- استخراج ویژگی با شبکه پایه: در مرحله اول، جهت استخراج ویژگی‌های اصلی از معماری VGG16 که بر روی مجموعه داده ImageNet از پیش آموزش‌دیده، استفاده می‌شود. لایه‌های کانولوشنی این معماری توانایی بالایی در استخراج ویژگی‌های اصلی و سلسله‌مراتبی از تصاویر ورودی دارند. وردی این بخش تصاویر ورق‌های فولاد و خروجی این بخش، یک نقشه ویژگی چندکاناله است که به ماژول توجه منتقل می‌شود.

۲- ماژول توجه: در این مرحله خروجی مرحله اول شامل ویژگی‌های چندکاناله به شبکه عصبی کانولوشن وارد شده و اعمال زیر به‌ترتیب روی آن اجرا می‌شود:

تولید نگاشت توجه: ایجاد یک نگاشت توجه فضایی تک‌کاناله با استفاده از تابع فعال‌ساز Sigmoid که اهمیت هر ناحیه مکانی را با مقداری بین ۰ و ۱ مشخص می‌کند.

ماسک کردن ویژگی‌ها: استفاده از نگاشت توجه تولیدشده به‌عنوان یک ماسک که به‌صورت تک‌به‌تک در نقشه ویژگی‌های اصلی ضرب می‌شود. این کار باعث کم‌رنگ شدن اهمیت ویژگی‌های موجود در پس‌زمینه تصاویر فولاد و تقویت ویژگی‌های ناحیه اصلی تصاویر می‌شود و خروجی ویژگی‌های ماسک شده است.

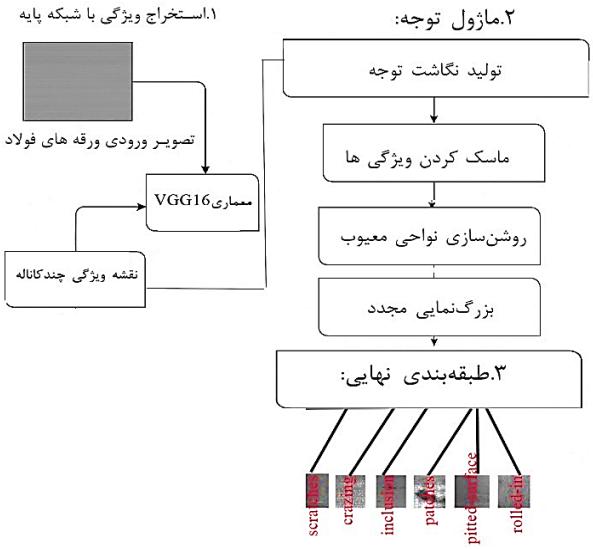
روشن‌سازی نواحی معیوب: در این مرحله ویژگی‌های ماسک شده، جهت روشن‌سازی نواحی معیوب و کاهش ابعاد وارد لایه‌های میانگین‌گیری سراسری (GAP[3])، می‌شوند. خروجی این بخش به بخش بزرگ‌نمایی مجدد جهت نرمال‌سازی تطبیقی وارد می‌شود.

بزرگ‌نمایی مجدد: در این بخش، خروجی لایه قبل به یک لایه لامبدا[4] وارد می‌شود که طی آن، میانگین ویژگی‌های ماسک شده بر میانگین نگاشت توجه تقسیم می‌گردد. این فرایند مانند یک نرمال‌سازی تطبیقی عمل کرده که با توجه به امکان پیاده‌سازی آن در چارچوب Keras و TensorFlow امکان کاهش اثر پراکندگی در نقشه توجه (در مدل پیشنهادی) را افزایش می‌دهد. همین امر از حذف الگوهای ظریف و عیوب کوچک سطوح فولاد جلوگیری می‌کند. بدین‌ترتیب، حساسیت مدل در تشخیص جزئیات بسیار کوچک افزایش‌یافته و عملکرد کلی شبکه بهبود قابل‌توجهی پیدا می‌نماید. در انتهای این بخش یک لایه Dropout جهت جلوگیری از بیش برازش قرار می‌گیرد و خروجی به بخش طبقه‌بندی وارد می‌شود.

۳- طبقه‌بندی نهایی: در بخش پایانی، خروجی حاصل از بخش بزرگ‌نمایی مجدد به یک شبکه متصل کامل وارد شده و با استفاده از تابع فعال‌ساز Softmax، تصاویر در ۶ کلاس ترک‌خوردگی شبکه‌ای[5]، شامل ناخالصی[6]، لکه‌ها[7]، سطح حفره‌دار[8]، پوسته نوردی[9]، خراش‌ها[10] با توجه به نوع خرابی، طبقه‌بندی می‌شوند.

[1] Legendre Multiwavelet Transform
[2] Feature Attention Guidance
[3] Global Average Pooling
[4] Lambda
[5] Crazing
[6] Inclusion
[7] Patches
[8] Pitted Surface
[9] Rolled-In Scale
[10] Scratches

در شکل ۱ فلوچارت مدل پیشنهادی نمایش داده شده است.



شکل (۱): فلوچارت مدل پیشنهادی

۳-۲ شرح لایه‌های مدل

لایه کانولوشن[1]: این لایه بلوک اصلی سازنده شبکه عصبی کانولوشن است. در این لایه، یک فیلتر بر روی تصویر ورودی به‌صورت گام‌به‌گام حرکت کرده و با انجام عملیات کانولوشن (ضرب ماتریسی)، ویژگی‌های محلی مانند لبه‌ها و بافت‌ها را در قالب نقشه‌های ویژگی استخراج می‌کند. خروجی این لایه از طریق رابطه ۱ محاسبه می‌شود:

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i.j}^{(l-1)} \times C_J^{(l)} \qquad (۱)$$

در این مدل، Bi(l) ماتریس بایاس در لایه $\boldsymbol{l}$ است. فیلتر یا هسته کانولوشن با نماد $K_{i.j}^{(l-1)}$ نشان داده می‌شود که m*m اندازه آن است و برای اتصال نگاشت ویژگی شماره j در لایه l-1 به نگاشت ویژگی شماره i در لایه $\boldsymbol{l}$ به کار می‌رود. خروجی لایه کانولوشن با نماد $C_i^{(l)}$ نمایش داده می‌شود. ورودی لایه کانولوشن اولیه با نماد Ci(l − 1) مشخص می‌شود که در اولین لایه کانولوشن برابر با تصویر ورودی ورق‌های فولاد و Xi است.

توابع فعال‌سازی[2]: پس از هر لایه کانولوشن، از یک تابع فعال‌سازی برای غیرخطی بودن مدل استفاده می‌شود. در این تحقیق از توابع زیر استفاده‌شده است:

تابع فعال‌ساز Sigmoid: مقادیر ورودی را به بازه [۰, ۱] نگاشت می‌دهد و برای تولید نگاشت توجه ایده‌آل است.

تابع فعال‌ساز ELU[3]: با جلوگیری از مشکل نورون‌های مرده، به همگرایی سریع‌تر و دقیق‌تر مدل کمک می‌کند.

خطی[4]: یک تبدیل خطی ساده را بدون اعمال غیرخطی بودن انجام می‌دهد.

لایه ادغام میانگین[5]: این لایه با محاسبه میانگین مقادیر در یک ناحیه مشخص، ابعاد فضایی نقشه‌های ویژگی را کاهش می‌دهد.

لایه متصل‌کننده کامل[6]: این لایه که در انتهای مدل قرار دارد، وظیفه کلاس‌بندی (طبقه‌بندی) را برعهده دارد.

لایه نرمال‌ساز دسته‌ای[7]: این لایه با نرمال‌سازی خروجی‌های لایه قبلی در هر دسته، باعث پایداری و آموزش سریع‌تر مدل می‌شود.

لایه ادغام میانگین جهانی (GAP): یک بخش ساختاری است که در انتهای بخش استخراج ویژگی به کار می‌رود. وظیفه اصلی آن، تقلیل ابعاد فضایی[8] نگاشت‌های ویژگی است. این کار را با محاسبه میانگین کل مقادیر در هر کانال ویژگی انجام می‌دهد و حاصل تغییر هر نگاشت ویژگی یک مقدار اسکالر واحد است. این عملیات به‌عنوان یک تنظیم‌کننده[9] مؤثر، از پدیده بیش برازش[10] جلوگیری کرده و قابلیت تعمیم مدل را به شکل قابل‌توجهی بهبود می‌بخشد [۱۶].

در مدل پیشنهادی، برای بهره‌برداری از مزایای GAP در کنار حفظ حداقلی اطلاعات مکانی، لایه به‌صورت «ادغام میانگین هم‌پوشان[11]» با گام[12] برابر با (۱، ۱) پیاده‌سازی شده است. بررسی مدل با حذف لایه GAP از ساختار شبکه، منجر به نوسانات جزئی در منحنی

[1] Convolutional Layer
[2] Activation Functions
[3] Exponential Linear Unit
[4] Linear
[5] Average Pooling Layer
[6] Fully Connected Layer
[7] Batch Normalization
[8] Spatial Dimension
[9] Regularizer
[10] Overfitting
[11] Overlapping Average Pooling
[12] Stride

تابع زیان نهایی[1] در طول فرایند آموزش شد. این مشاهده تجربی، نقش حیاتی لایه GAP در تثبیت فرایند آموزش و تضمین قابلیت اطمینان مدل در طبقه‌بندی عیوب سطح ورق‌های فولاد را به اثبات رساند.

لایه لامبدا: این لایه در مدل‌های یادگیری عمیق به‌عنوان ابزاری توانمند به‌منظور تعریف عملیات ریاضی سفارشی درون گراف محاسباتی شبکه عصبی به کار می‌رود. این لایه امکان اجرای هر نوع تابع دل‌خواه $\boldsymbol{f(0)}$را فراهم می‌کند، بدون آنکه نیاز به تعریف یک کلاس جدید باشد و در عین حال از طریق پس‌انتشار، قابلیت تفاضل‌پذیری خود را برای آموزش حفظ می‌کند [۱۶].

در این پژوهش، از لایه لامبدا به‌منظور پیاده‌سازی فرایند بزرگ‌نمایی مجدد تطبیقی مبتنی بر توجه استفاده‌شده است. نتیجه آن، افزایش حساسیت مدل نسبت به عیوب سطحی کوچک در تصاویر ورق‌های فولاد است. این لایه پس از محاسبه میانگین ویژگی‌های ماسک شده و میانگین نگاشت توجه، نسبت این دو مقدار را به‌عنوان یک نرمال‌سازی تطبیقی محاسبه می‌کند.

در ادامه مراحل محاسباتی مدل شامل استخراج ویژگی و ماژول توجه و طبقه‌بندی به‌صورت مرحله‌به‌مرحله بیان شده است. در ابتدا نمادهای محاسباتی در جدول ۱ تشریح شده است.

جدول (۱): نمادهای محاسباتی مدل پیشنهادی

نماد	توضیح
F_bn	ویژگی‌های استخراج شده از شبکه پایه VGG16 پس از نرمال‌سازی
H,W	ارتفاع و عرض نقشه ویژگی‌ها، ابعاد مکانی تصویر (200, 200)
C	تعداد کانال‌های ویژگی [۳]
A	نگاشت توجه
F_masked	ویژگی‌های ماسک شده
GAP	میانگین سراسری ویژگی‌های ماسک شده
$\bar{A}$	میانگین کل مقادیر نگاشت توجه
ε	مقدار کوچک (برای جلوگیری از تقسیم بر صفر)
$v_rescaled$	ویژگی‌های نهایی پس از بزرگ‌نمایی تطبیقی
F_vgg	خروجی استخراج ویژگی
σ(·)	تابع فعال‌ساز Sigmoid
$Conv1 \times 1$	لایه کانولوشن با هسته ۱×۱ برای فشرده‌سازی کانال‌ها
$f(0)$	تابع فعال‌ساز غیرخطی (مانند ELU یا ReLU)
$AvgPool2 \times 2(0)$	میانگین‌گیری محلی برای صاف‌سازی مکانی
W_1, b_1	وزن و بایاس لایه Dense اول در طبقه‌بندی
W_2, b_2	وزن و بایاس لایه Dense دوم در طبقه‌بندی
h_1	خروجی لایه Dense اول با فعال‌ساز ReLU
y	بردار احتمال طبقه‌بندی خروجی Softmax
$\hat{c}$	کلاس با طبقه خرابی پیش‌بینی شده مدل ٦

۱- ورودی تصویر در رابطه ۲ نمایش داده شده است:

$$I \in \mathbb{R}\hat{}(H \times W \times C) \quad (۲)$$

۲- استخراج ویژگی‌ها با **VGG16** و نرمال‌سازی در رابطه ۳ و ٤ بیان شده است:

$$F_vgg = VGG16(I) \quad (۳)$$

$$F_bn = BatchNorm\big(MaxPool(F_vgg)\big), \quad (٤)$$

۳- تولید نگاشت توجه در رابطه ۵ بیان شده است:

$$A = \sigma(Conv1 \times 1(AvgPool2 \times 2(f(Conv1 \times 1(Fbn))))) \quad (۵)$$

٤- ماسک‌گذاری ویژگی‌ها در رابطه ٦ بیان شده است:

$$F_masked = A \odot F_bn, \quad (٦)$$

۵- کاهش ابعاد با لایه ادغام میانگین جهانی در رابطه ۷ بیان شده است:

$$F\_masked = GAP(F\_masked)$$
$$\bar{A} = GAP(A) \quad (۷)$$

٦- نرمال‌سازی تطبیقی با لایه لامبدا در رابطه ۸ بیان شده است:

$$v_rescaled(k) = F_masked(k)/(\bar{A} + \varepsilon) \quad (۸)$$

۷- طبقه‌بندی در رابطه ۹ بیان شده است:

[1] Loss

$$y = Softmax\big(W_2 \cdot ReLU(W_1(v_r escaled(k)) + b_1) + b_2\big) \quad (۹)$$
$$\hat{c} = arg\,\max_i(y_i)$$

۳-۳ مجموعه داده

در این تحقیق از دو مجموعه داده معتبر NEU-DT [۱۷]و NEU-CLS [۱۸] برای طبقه‌بندی خرابی‌های سطحی در نوارهای فولادی نورد گرم که توسط دانشگاه Northeastern ایجاد شده است، استفاده‌شده است. ویژگی‌های کلی دو مجموعه داده به شرح زیر است:

تعداد تصاویر: ۱۸۰۰ تصویر خاکستری[1]، با ابعاد ۲۰۰×۲۰۰ پیکسل

طبقه‌ها (۶ نوع عیب):

با توجه به انواع خرابی شامل ترک‌خوردگی سطحی[2]، درون ماندگی[3]، پچ‌ها[4]، سطح حفره‌دار[5]، پوسته نوردی[6] و خراش‌ها[7] می‌شود و تعداد نمونه در هر کلاس ۳۰۰ تصویر است.

۱- مجموعه داده NEU-DT

این مجموعه داده برای تشخیص اشیا به کار می‌رود و از تعداد ۱۸۰۰ تصویر به نسبت ۸۰ به ۲۰ در دو پوشه Train و Validiation قرار داد و حاوی یک فایل برچسب‌ها[8] به شرح زیر است:

فایل‌های XML حاوی برچسب کلاس و مختصات برای هر خرابی مناسب برای عملیات تشخیص اشیا است و برای هر تصویر در پوشه‌های Train و Validiation یک فایل XML وجود دارد. در انتها ۱۴۴۸ تصویر در پوشه Train و ۳۶۰ تصویر در پوشه Validiation قرار دارد.

۲- مجموعه داده NEU-CLS

این مجموعه داده برای طبقه‌بندی تصاویر به کار می‌رود و تعداد ۱۸۰۰ تصویر، ۱۷۷۰ تصویر در پوشه Train و ۳۰ تصویر در پوشه Validiation قرار داد و فقط دارای برچسب کلاس (نوع عیب) بوده و مختصات محل عیب ارائه نمی‌شود.

۴-۳ ساختار الگوریتم مدل پیشنهادی در پایتون

۱- استخراج ویژگی با شبکه پایه:

در ابتدا پیش‌پردازش انجام می‌شود و اندازه تصاویر از سایز ۲۰۰×۲۰۰×۱ به سایز ۲۰۰×۲۰۰×۳ تغییر پیدا می‌کند و وارد لایه ورودی معماری VGG16 با یک لایه فریز شده می‌شوند و تحت پیش‌آموزش قرار می‌گیرند که سبب افزایش تعداد پارامترها و اتصالات از ۰ به ۲/۳۵۹/۸۰۸ و استخراج ویژگی‌های شبکه پایه به‌صورت نقشه ویژگی چندکاناله می‌شود.

۲- ماژول توجه:

خروجی نقشه ویژگی مرحله قبل وارد ماژول توجه می‌شود و مراحل زیر را به‌ترتیب طی می‌نماید:

- تولید نگاشت توجه

این فرایند شامل چندلایه کانولوشنی متوالی با فیلترهای ۱×۱ است و هدف آن علاوه‌بر فشرده‌سازی تدریجی، استخراج اطلاعات مکانی و ویژگی مهم است:

لایه توجه اول: ویژگی‌های استخراج شده از مرحله قبل وارد یک لایه کانولوشن با ۱۲۸ فیلتر در اندازه ۱×۱ می‌شوند و پس از عمل ضرب کانولوشنی و اعمال تابع فعال‌ساز Elu به لایه توجه دوم منتقل می‌شود. این لایه سبب افزایش تعداد پارامترها و اتصالات از ۲۰۴۸ به ۶۵۶۶۴ می‌گردد.

لایه توجه دوم: خروجی لایه توجه اول به لایه توجه دوم که یک لایه کانولوشن با ۳۲ فیلتر در اندازه ۱×۱ است، وارد می‌شود و پس از عمل ضرب کانولوشنی و اعمال تابع فعال‌ساز Elu به لایه توجه دوم منتقل می‌شود. این لایه سبب کاهش تعداد پارامترها و اتصالات از ۶۵۶۶۴ به ۴۱۲۸ می‌گردد.

لایه توجه سوم: خروجی لایه توجه دوم به لایه توجه سوم که یک لایه کانولوشن با ۱۶ فیلتر در اندازه ۱×۱ است، وارد می‌شود و پس از عمل ضرب کانولوشنی و اعمال تابع فعال‌ساز Elu به لایه ادغام متوسط منتقل می‌شود. این لایه سبب کاهش تعداد پارامترها و اتصالات از ۴۱۲۸ به ۵۲۸ می‌گردد.

[1] Grayscale
[2] Crazing
[3] Inclusion
[4] Patches
[5] Pitted Surface
[6] Rolled-in Scale
[7] Scratches
[8] Annotations

لایه ادغام متوسط: خروجی لایه توجه سوم به لایه ادغام متوسط وارد می‌شود که پنجره ۲×۲ دارد و داده‌ها به میزان متوسط کاهش می‌یابد. در انتهای این لایه تعداد پارامترها به ۰ می‌رسد.

لایه توجه چهارم: خروجی لایه ادغام متوسط به لایه توجه چهارم که یک لایه کانولوشن با ۱۶ فیلتر در اندازه ۱×۱ است، وارد می‌شود و پس از عمل ضرب کانولوشنی و اعمال تابع فعال‌ساز Sigmoid، تعداد پارامترها و اتصالات از ۰ به ۱۷ تغییر پیدا می‌کند.

- ماسک کردن ویژگی‌ها

در این بخش، نگاشت توجه تولیدشده به‌عنوان ماسک بر ویژگی‌های اصلی اعمال می‌شود:

لایه کانولوشن: یک لایه کانولوشن با ۵۱۲ فیلتر در اندازه ۱×۱ در نگاشت توجه ضرب ماتریسی می‌شود و پس از اعمال تابع فعال‌ساز خطی به لایه چندمنظوره منتقل می‌شود. این لایه سبب افزایش تعداد پارامترها و اتصالات از ۱۷ به ۵۱۲ می‌گردد.

لایه چندمنظوره: خروجی لایه کانولوشن قبلی برای ماسک کردن ویژگی‌ها به لایه چندمنظوره وارد می‌شود (لایه چندمنظوره به لایه نرمال‌ساز اول و لایه کانولوشن در این بخش متصل است). در پایان این مرحله تعداد پارامترها و اتصالات از ۵۱۲ به ۰ کاهش می‌یابد.

- روشن‌سازی نواحی معیوب و کاهش ابعاد

در این بخش دو لایه ادغام میانگین جهانی در جهت روشن‌سازی نواحی معیوب قرار دارد. لایه اول به لایه چندمنظوره و لایه دوم ادغام میانگین جهانی به لایه کانولوشن این بخش است.

- بزرگ‌نمایی مجدد

خروجی مرحله قبل به یک لایه لامبدا وارد می‌شود تا علاوه‌بر کاهش اندازه، فرایند نرمال‌سازی تطبیقی هم انجام شود. در نهایت، یک لایه Dropout به‌منظور جلوگیری از بیش‌برازش[1] به مدل اضافه می‌شود.

- طبقه‌بندی نهایی

در این بخش، دو لایه متصل‌کننده کامل جهت طبقه‌بندی به مدل اضافه می‌شود:

لایه متصل‌کننده کامل اول:

خروجی مرحله بزرگ‌نمایی مجدد به لایه متصل‌کننده کامل اول با ۱۲۸ نورون و تابع فعال‌ساز Elu وارد می‌شود. سپس یک لایه Dropout به‌منظور جلوگیری از بیش‌برازش به این لایه متصل می‌شود.

جدول (۲): ساختار مدل پیشنهادی در پایتون

مدل	لایه‌ها	سایز داده‌ها	پارامترها
استخراج ویژگی	input_layer_1	None, 200,) (200, 3	۰
	VGG16-1layer	None, 12, 12,) (512	۲/۳۵۹/۸۰۸
	MaxPooling2D	None, 6, 6,) (512	۰
	BhNormalization	None, 6, 6,) (512	۲۰۴۸
تولید نگاشت توجه	conv2d-1	None, 6, 6,) (128	٦٥٦٦٤
	conv2d-2	None, 6, 6,) (32	٤١٢٨
	conv2d-3	None, 6, 6,) (16	٥٢٨
	AveragePooling2D	None, 6, 6,) (16	۰
	conv2d-4	(None, 6, 6, 1)	۱۷
ماسک	conv2d	None, 6, 6,) (512	٥١٢
	multiply	None, 6, 6,) (512	۰
کاهش ابعاد	GlobalAePooling-1	(None, 512)	۰
	GlobalAePooling-2	(None, 512)	۰
بزرگ‌نمایی	Lambda	(None, 512)	۰
	dropout	(None, 512)	۰
طبقه‌بندی	dense_1	(None, 128)	٦٥/٦٦٤
	dropout_2	(None, 128)	۰
	dense_2	(None, 6)	٧٧٤

لایه متصل‌کننده کامل دوم:

خروجی لایه قبلی جهت کلاس‌بندی به این لایه با ٦ نورون و تابع فعال‌ساز Softmax وارد می‌شود و برحسب تعداد عیوب موجود در مجموعه داده به ٦ کلاس ترک‌خوردگی شبکه‌ای، شامل ناخالصی، لکه‌ها، سطح حفره‌دار، پوسته نوردی و خراش‌ها طبقه‌بندی می‌شوند. در جدول ۲ هر ۵ بخش مدل پیشنهادی شامل

[1] Overfitting

استخراج ویژگی، تولید نگاشت توجه، ماسک، کاهش ابعاد و طبقه‌بندی به‌همراه نام لایه‌ها، سایز داده‌ها در لایه‌ها و تعداد پارامترهای حاصل از هر لایه در شبیه‌ساز پایتون نمایش داده شده است. کد در [۱۹] در دسترس است.

۳-۵ روش ارزیابی

برای ارزیابی طبقه‌بندی انواع خرابی‌ها در تصاویر فولاد از معیارهای امتیاز F1[1]، بازخوانی[2]، دقت و صحت[3] استفاده می‌شود. در معیارهای بالا ٤ پارامتر زیر استفاده می‌شود:

TP[4]: پیش‌بینی درست برای نمونه‌های مثبت (مثلاً ورق فولاد خراب را خراب طبقه‌بندی داده).

FP[5]: پیش‌بینی اشتباه که مدل نمونه منفی را مثبت طبقه‌بندی داده (مثلاً ورق فولاد سالم را خراب طبقه‌بندی داده).

FN[6]: پیش‌بینی اشتباه که مدل نمونه مثبت را منفی طبقه‌بندی داده (مثلاً ورق فولاد خراب را سالم اعلام کرده).

TN[7]: پیش‌بینی درست برای نمونه‌های منفی (مثلاً ورق فولاد سالم را سالم طبقه‌بندی داده).

دقت معیاری است که حاصل از تقسیم تعداد موارد درست طبقه‌بندی داده شده به کل موارد است. دقت، مجاورت یک مقدار محاسبه‌شده با یک مقدار طبیعی یا واقعی است. به عبارت دیگر، این معیار ارزیابی می‌تواند مقدار دقیقی که دقت آن قابل‌اندازه‌گیری است را اندازه‌گیری کند، از رابطه ۱۰ محاسبه می‌شود:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad (۱۰)$$

دقت پیش‌بینی مثبت‌ها: برای یک کلاس تعداد موارد واقعاً طبقه‌بندی‌شده (دارای برچسب صحیح) تقسیم بر جمع اقلام درست یا نادرست با برچسب متعلق به آن طبقه است، از رابطه ۱۱ محاسبه می‌شود:

$$\text{precision} = \frac{TP}{TP+FP} \qquad (۱۱)$$

بازخوانی: اشاره به کسری از مواردی که به‌درستی طبقه‌بندی‌شده به تعداد کل موارد طبقه‌بندی‌شده در یک کلاس، از رابطه ۱۲ محاسبه می‌شود:

$$\text{recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (۱۲)$$

امتیاز F1: معیاری ترکیبی برای صحت و بازخوانی. مخصوصاً در داده‌های نامتوازن کاربرد زیادی دارد، از رابطه ۱۳ محاسبه می‌شود [۲۰].

$$\text{F1\_Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (۱۳)$$

٤- شبیه‌سازی و ارزیابی

با توجه به توضیحات ارائه‌شده در بخش‌های قبل، در این مقاله از ۱٤٤۸ تصویر پوشه Train از مجموعه داده NEU-DET و NEU-CLS شامل انواع خرابی‌های سطحی در نوارهای فولادی نورد گرم، استفاده‌شده است. نتایج با اجرای شبیه‌سازی در محیط اسپایدر در نسخه ۵ و در لپ‌تاپ با پردازنده Corei7 و حافظه تصادفی ۲۰ گیگابایت به‌دست‌آمده است. نتایج این شبیه‌سازی مرحله‌به‌مرحله به نمایش گذاشته می‌شود:

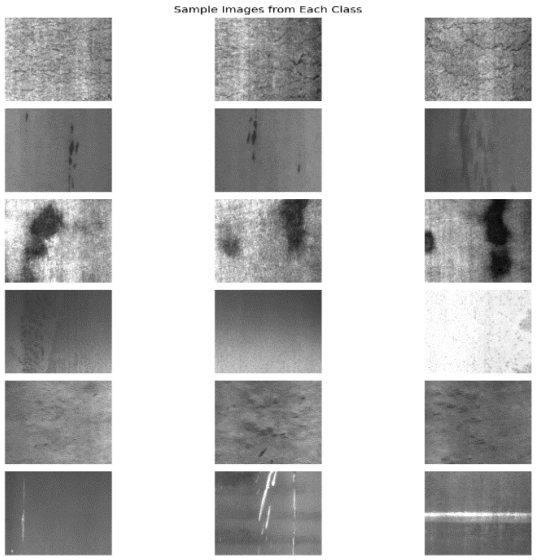
٤-۱ اجرای آنالیز اکتشافی داده‌ها

در این مرحله آنالیز اکتشافی داده‌ها بر ۱٤٤۸ تصویر پوشه Train اجرا شده است. در شکل ۲ تعداد ۱۸ تصویر به‌صورت تصادفی از کلاس‌های (طبقه‌های) مختلف نمایش داده شده است.

٤-۲ اجرای مدل

در این مرحله مدل پیشنهادی بر روی داده‌های Train طبق جدول ۳ که حاوی ابر پارامترهاست، اجرا می‌شود.

[1] F1-Score
[2] Recall
[3] Precision
[4] True Positive
[5] False Positive
[6] False Negative
[7] True Negative

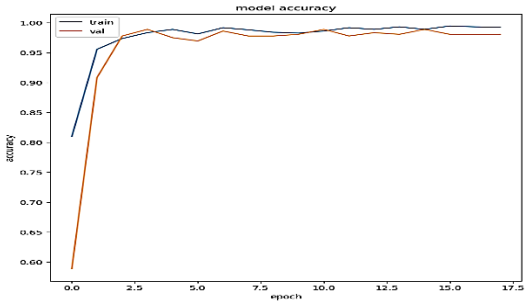


شکل (۲): تعداد ۱۸ تصویر از کلاس‌های مختلف در Train

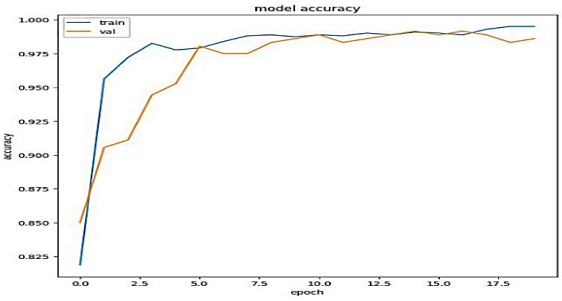
جدول (۳): ابر پارامترهای مدل پیشنهادی

توضیحات	مقدار/انتخاب	پارامتر
سایز تصویر ورودی	۲۰۰*۲۰۰*۳	ورودی مدل
اندازه بسته‌ها	٦٤	اندازه Batch
تعداد دور آموزش	٥٠	تعداد epoch
نرخ لایه‌ها	(۰/۵، ۰/۲۵)	نرخ Dropout
تابع فعال‌ساز GAP	igmoidS ،ELU	تابع فعال‌ساز
بهینه‌ساز مدل	Adam	بهینه‌ساز
تابع هزینه برای ٦ کلاس	Categorical Cross-Entropy	تابع هزینه
توقف زودهنگام	Stop_Callback	Callback
تعداد تصویر	۱۱۵۸ تصویر	Train
تعداد تصویر	۲۹۰ تصویر	val
تعداد تصویر	۳٦۰ تصویر	Validiation
نرخ آموزش	۰/۰۰۰۱	Learning Rate

در ادامه نمودار دقت مدل پیشنهادی (این شکل روند تغییر دقت مدل را در طول فرایند آموزش نمایش می‌دهد. محور افقی نشان‌دهنده شماره دوره‌های آموزشی و محور عمودی مقدار دقت مدل است. دو منحنی به‌ترتیب آبی و قرمز، عملکرد مدل روی داده‌های آموزش و ارزیابی را نشان می‌دهند.) در مجموعه داده NEU-DET در شکل ۳ و در مجموعه داده NEU-CLS در شکل ٤ نمایش داده شده است:



شکل (۳): نمودار دقت در مجموعه داده NEU-DET



شکل (٤): نمودار دقت در مجموعه داده NEU-CLS

٤-۳ ارزیابی

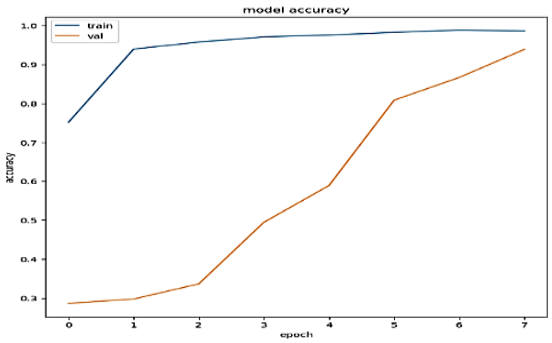
ارزیابی در دو سناریو انجام می‌شود:

سناریوی اول: در این سناریو مدل پیشنهادی بدون بخش ماژول توجه و فقط با دو بخش استخراج ویژگی‌ها با VGG16 و طبقه‌بندی بر روی داده‌های بخش Test و Train اجرا می‌شود. در جدول ٤ بخش‌های مدل در سناریوی ۲ به‌همراه نام لایه‌ها، سایز داده‌ها در لایه‌ها و تعداد پارامترهای حاصل از هر لایه در شبیه‌ساز پایتون نمایش داده شده است.

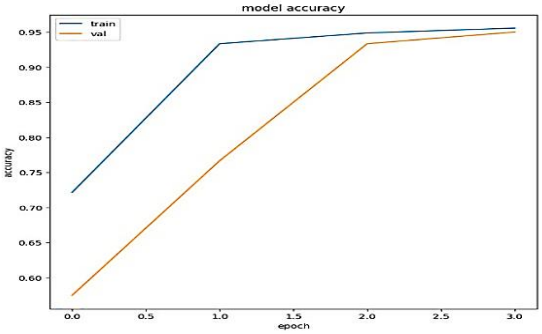
برای اجرای مدل بدون ماژول توجه، از جدول ابر پارامترهای ۳ استفاده‌شده است. نمودار دقت سناریوی ۲ (این شکل روند تغییر دقت مدل را در طول فرایند آموزش نمایش می‌دهد. محور افقی نشان‌دهنده شماره دوره‌های آموزشی و محور عمودی مقدار دقت مدل است. دو منحنی به‌ترتیب آبی و قرمز، عملکرد مدل روی داده‌های آموزش و ارزیابی را نشان می‌دهند.) در مجموعه داده NEU-DET در شکل ۵ و در مجموعه داده NEU-CLS در شکل ٦ نمایش داده شده است:

جدول (٤): ساختار جدولی مدل بدون ماژول توجه در پایتون

پارامترها	سایز داده‌ها	لایه‌ها	مدل
۰	(None, ۲۰۰, ۲۰۰, 3)	input_layer_1	استخراج ویژگی
۱٤/۷۱۵/۷۱۲	(None, 6, 6, 512)	VGG16--1layer	
۰	(None, 512)	global_average_pooling2d	
۰	(None, 512)	dropout_2	
٦٥/٦٦٤	(None, 128)	dense_2	طبقه‌بندی
۰	(None, 128)	dropout_3	
۷۷٤	(None, ٦)	dense_3	



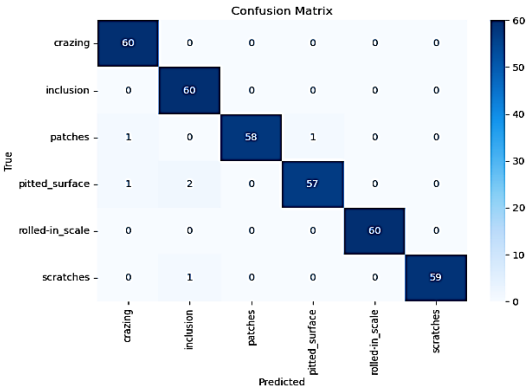
شکل (٥): نمودار دقت مدل در سناریوی ۱ در NEU-DET



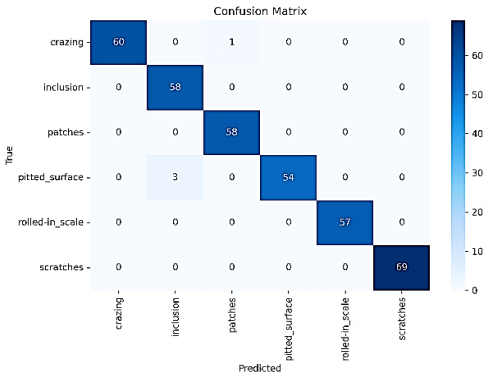
شکل (٦): نمودار دقت مدل در سناریوی ۱ در NEU-CLS

در انتها در شکل ۷ و ۸ ماتریس درهم‌ریختگی (در این ماتریس، محور افقی کلاس‌های پیش‌بینی شده و محور عمودی کلاس‌های واقعی قرارگرفته‌اند. در هر سلول، تعداد نمونه‌هایی که کلاس واقعی‌شان در سطر موردنظر بوده و مدل پیش‌بینی کرده و متعلق به کلاس ستون متناظر هستند، نمایش داده شده است) در مجموعه داده NEU-DET و در مجموعه داده NEU-CLS نمایش داده شده است:

در جدول ۵ نتایج ارزیابی مدل را در دو سناریو و برای دو مجموعه داده مختلف (NEU-DET و NEU-CLS) نمایش می‌دهد. هر سناریو با مجموعه‌ای از معیارهای ارزیابی اصلی مدل طبقه‌بندی مانند دقت، دقت مثبت، بازخوانی و امتیاز F1 بررسی‌شده است:



شکل (۷): ماتریس درهم‌ریختگی سناریوی ۱ در NEU-DET



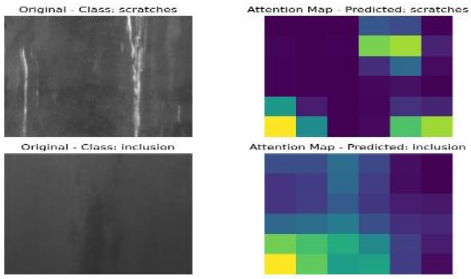
شکل (۸): ماتریس درهم‌ریختگی سناریوی ۱ NEU-CLS

جدول (٥): نتایج سناریوی اول (در قالب درصد)

روش	دقت	دقت مثبت	بازخوانی	امتیاز F1	مجموعه داده
سناریو ۱	۰/۹٤	۹٤/٦۱	۹۳/۸۹	۹٤/۲۱	NEU-DET
سناریو ۱	۹٦/۹٥	۹٥	۹٥/۰۸	۹٤/۹٤	NEU-CLS

سناریوی دوم: در این سناریو مدل پیشنهادی بر روی ۳٦۰ تصویر به‌عنوان Validiation اجرا شده و در شکل ۹ و ۱۰ نتایج حاصل از کلاس‌بندی و پیش‌بینی نوع خرابی در دیتاست NEU-DET و NEU-CLS، در شکل ۱۱ و ۱۲ ماتریس درهم‌ریختگی (در این ماتریس، محور افقی کلاس‌های پیش‌بینی شده و محور عمودی

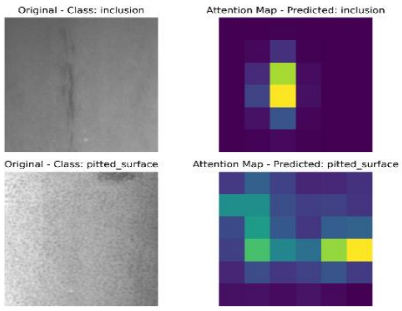
کلاس‌های واقعی قرارگرفته‌اند) در مجموعه داده NEU-DET و مجموعه داده NEU-CLS و جدول ٦ نتایج ارزیابی مدل را در دو سناریو و برای دو مجموعه داده مختلف (NEU-DET و NEU-CLS) نمایش می‌دهد. هر سناریو با مجموعه‌ای از معیارهای ارزیابی اصلی مدل طبقه‌بندی مانند دقت، دقت مثبت، بازخوانی و امتیاز F1 بررسی‌شده است:



شکل (۹): نتایج حاصل از کلاس‌بندی خرابی در NEU-DET

شکل ۹ عملکرد مدل و نقشه توجه مدل را برای دو نمونه تصویر از دیتاست NEU-DET نشان می‌دهد:

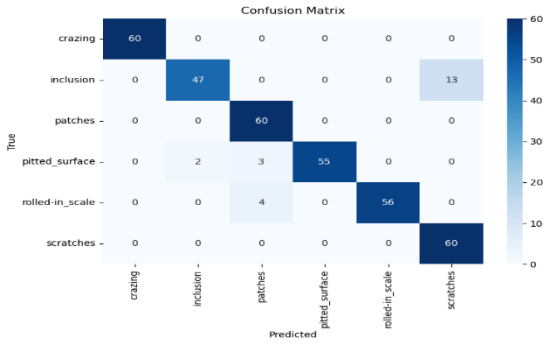
- سمت چپ، تصویر واقعی با برچسب کلاس Scratches (بالا) و Inclusion (پایین) قرار دارد.
- سمت راست، نقشه توجه مدل پس از طبقه‌بندی موفق هر نمونه مشاهده می‌شود.



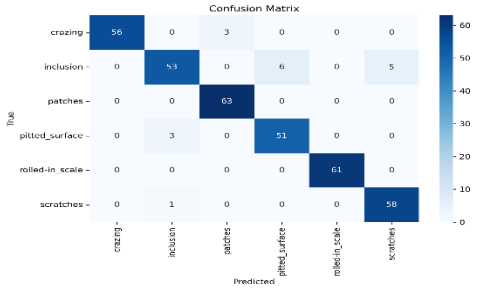
شکل (۱۰): نتایج کلاس‌بندی خرابی در NEU-CLS

شکل ۱۰ عملکرد مدل و نقشه توجه مدل را برای دو نمونه تصویر از دیتاست **NEU-CLS** نشان می‌دهد:

- سمت چپ، تصویر واقعی با برچسب کلاس Inclusion **(بالا)** و Pitted_Surface (پایین) قرار دارد.
- سمت راست، نقشه توجه مدل پس از طبقه‌بندی موفق هر نمونه مشاهده می‌شود.



شکل (۱۱): ماتریس درهم‌ریختگی سناریوی ۲ در NEU-DET



شکل (۱۲): جدول درهم‌ریختگی سناریوی ۲ در NEU-CLS

جدول (٦): نتایج سناریو دوم (در قالب درصد)

روش	دقت	دقت مثبت	بازخوانی	امتیاز F1	مجموعه داده
سناریو ۲	۹۹/۹۲	۹۸/۳۸	۹۸/۳۳	۹۹	NEU-DET
سناریو ۲	۹۹/۹۸	۹۸/۹۰	۹۸/۸۵	۹۹	NEU-CLS

۵- ساخت وب اپلیکیشن طبقه‌بندی

در این مقاله، جهت طبقه‌بندی خودکار انواع خرابی‌ها در ورق‌های فولادی یک وب اپلیکیشن کاربردی با استفاده از فریم‌ورک Streamlit ساخته‌شده است. این ابزار، به کاربران امکان می‌دهد تا با بارگذاری تصاویر ورق‌های فولادی، نوع عیوب موجود را مشخص کند. این اپلیکیشن به‌صورت خصوصی در [۲۱] دسترس است و در شکل ۱۳، تصویری از رابط کاربری آن نمایش داده شده است.

شکل (۱۳): وب اپلیکیشن طبقه‌بندی

شکل ۱۳ تصویری نمونه از سامانه تشخیص عیوب سطح فولاد مبتنی بر مدل VGG16 و نقشه توجه (Grad-CAM) را نمایش می‌دهد. توضیحات اجزای شکل:

۱- کاربر می‌تواند فایل تصویر سطح فولاد را بارگذاری کند تا سامانه نوع عیب را تشخیص دهد.

۲- پس از بارگذاری، تصویر سطح فولاد به‌همراه نقشه توجه مدل نمایش داده می‌شود. نقشه توجه بخش‌هایی از تصویر را که مدل برای طبقه‌بندی مهم تشخیص داده، با رنگ‌های نموداری (قرمز، زرد، آبی) نشان می‌دهد.

در سمت راست، میزان احتمال مدل برای هر کلاس (طبقه) عیب به‌صورت عددی آورده شده است. در این نمونه، تشخیص مدل کلاس ترک‌خوردگی سطحی است که با دقت ۹۵/۴۸ درصد از سایر عیوب بسیار بالاتر است.

۶- مقایسه مدل‌های طبقه‌بندی عیوب سطح فولاد

در این بخش به مقایسه عملکرد مدل‌های طبقه‌بندی عیوب سطح فولاد براساس معیار دقت شامل، سناریوهای پیشنهادی در این مقاله و سایر تحقیقات در ۵ سال اخیر پرداخته می‌شود. نتایج در جدول ۷ بیان شده است.

همان‌طور که در جدول ۷ مشاهده می‌شود، سناریو ۲ با مدل‌های پیشین در دو رویکرد زیر برتری مطلق دارد:

مدل پیشنهادی در سناریو ۲ با دقت بالای ۹۹/۹۸ درصد نسبت به بهترین مدل موجود [۱۵] با ۹۹/۸۰ درصد بیش از ۰/۰۱۸٪ بهبود نشان داده است. این امر نمایانگر قدرت مدل و قدرت آن در تشخیص عیوب کوچک و سطحی ریز است.

جدول (۷): مقایسه مدل‌های طبقه‌بندی عیوب سطح فولاد

دقت[1]	مجموعه داده	مدل	محققان	سال
۹۹/۹۸	N[2]- CLS	Attention	سناریو ۱	۲۰۲۵
۹۵/۹۶	N-CLS	V[3]-Cnn	سناریو ۲	۲۰۲۵
۹۹/۰۶	N-CLS	LWT+ A-E	ژنگ وهمکاران	۲۰۲٤
۹۹/۸۰	N-CLS	LWT+FAG	ژنگ وهمکاران	۲۰۲۵
۹٤/۳۸	N-DET	V-Cnn	سناریو ۱	۲۰۲۵
۹۹/۹۲	N-DET	Attention	سناریو ۲	۲۰۲۵

۶-۱ دلایل اصلی قدرت مدل

۱- ماژول توجه: برجسته‌سازی نواحی معیوب و کاهش نویز.

۲- لایه لامبدا: نرمال‌سازی تطبیقی و کاهش ابعاد.

۳- **Dropout و Fully Connected**: جلوگیری از بیش‌برازش و بهبود تعمیم‌پذیری.

٤- **پایه VGG16**: استخراج ویژگی‌های عمیق و ترکیب مؤثر با توجه.

۷- نتیجه‌گیری و کارهای آینده

طبقه‌بندی دقیق عیوب سطح ورق‌های فولاد یکی از موضوعات کلیدی در ارتقای کیفیت و کاهش هزینه‌های تولیدات صنعتی است. بررسی پیشینه تحقیقات نشان می‌دهد اگرچه مدل‌های یادگیری عمیق اخیر مانند [۱۵] عملکرد خوبی ارائه کرده‌اند اما چالش اصلی که تمرکز بر عیوب کوچک در کاربردهای حساس است، همچنان برجاست.

در این پژوهش، جهت طبقه‌بندی عیوب موجود در ورق‌های فولاد، مدلی بهبودیافته از معماری VGG16 معرفی شد که با ادغام معماری قدرتمند VGG16 و ماژول توجه طراحی‌شده است. نتایج ارزیابی در دو مجموعه داده NEU-CLS و NEU-DET نشان دادند که این مدل در مقایسه با مدل سناریو ۱ (VGG16-CNN) و سایر رویکردهای موجود، بهبود چشمگیری در ٤ شاخص‌اصلی ارزیابی شامل دقت، پیش‌بینی مثبت، بازخوانی و امتیاز F1 داشته است و همگی به بالاتر از ۹۸٪ رسیده‌اند. نتایج به‌دست‌آمده

[1] Percent
[2] NEU
[3] VGG16

نشان‌دهنده قدرت مدل در برجسته‌سازی ویژگی‌های کلیدی، کاهش اثر نویز و افزایش تعمیم‌پذیری است.

بزرگ‌ترین نقطه‌ضعف مدل پیشنهادی علی‌رغم پیشرفت چشمگیر، ارزیابی مدل در محیط‌های واقعی صنعتی به‌شمار می‌رود. به‌طور کلی، نتایج این تحقیق نشان می‌دهد که ترکیب معماری‌های تثبیت شده مانند VGG16 با مکانیزم توجه، راهکاری نوین و کارآمد برای توسعه و تحلیل در سیستم‌های بازرسی هوشمند صنعت فولاد به‌شمار می‌رود.

کارهای آینده می‌تواند شامل چند مسیر متفاوت باشد. نخست، ارزیابی مدل بر روی مجموعه داده‌های متنوع‌تر و واقعی‌تر از عیوب سطحی ورق‌های فولاد، ضروری است. اگرچه مدل پیشنهادی بر روی دو مجموعه داده معتبر عملکرد مطلوبی نشان داد اما آزمودن آن بر روی داده‌های صنعتی حقیقی با شرایط نوری و نویز متفاوت، می‌تواند میزان پایداری و استحکام مدل را در شرایط واقعی مشخص سازد.

دوم، ترکیب مدل با سامانه‌های طبقه‌بندی و بازبینی خودکار کنترل کیفیت در صنعت فولاد است. این موضوع نه‌تنها سبب افزایش بهره‌وری می‌شود بلکه می‌تواند هزینه‌های ناشی از خطاهای انسانی را نیز کاهش دهد.

References

[1] K. Gong, D. Xu, F. Guo, Z. Wang, F. Zhang, and C. He, "Few-shot steel strip surface defect classification via self-correlation enhancement and feature refinement," IEEE Transactions on Instrumentation and Measurement, 2025.

[2] J. Huang, X. Zhang, L. Jia, and Y. Zhou, "An improved you only look once model for the multi-scale steel surface defect detection with multi-level alignment and cross-layer redistribution features," Engineering Applications of Artificial Intelligence, vol. 145, p. 110214, 2025.

[3] H. Chen, et al., "DCAM-Net: A rapid detection network for strip steel surface defects based on deformable convolution and attention mechanism," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-12, 2023.

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, Jun. 2017.

[5] Y. He, K. Song, Q. Meng, and Y. Yan, "An End-to-End Steel Surface Defect Detection Approach via Fusing Multiple Hierarchical Features," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 4, pp. 1493-1504, Apr. 2020.

[6] X. Cheng and J. Yu, "RetinaNet with Difference Channel Attention and Adaptively Spatial Feature Fusion for Steel Surface Defect Detection," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-11, 2021.

[7] R. Wei, Y. Song, and Y. Zhang, "Enhanced Faster Region Convolutional Neural Networks for Steel Surface Defect Detection," ISIJ International, vol. 60, no. 3, pp. 539-545, 2020.

[8] M. Tang, Y. He, J. Liu, K. Song, and Y. Yan, "A strip steel surface defect detection method based on attention mechanism and multi-scale maxpooling," Measurement Science and Technology, vol. 32, no. 11, p. 115401, 2021.

[9] Z. Guo, C. Wang, Y. Yang, Z. Wang, and F. Li, "Msft-yolo: Improved yolov5 based on transformer for detecting defects of steel surface," Sensors, vol. 22, no. 9, p. 3467, 2022.

[10] Z. Li, X. Wei, M. Hassaballah, Y. Li, and X. Jiang, "A deep learning model for steel surface defect detection," Complex & Intelligent Systems, vol. 10, no. 1, pp. 885-897, 2024.

[11] Y. Gao, G. Lv, D. Xiao, X. Han, T. Sun, and Z. Li, "Research on steel surface defect classification method based on deep learning," Scientific Reports, vol. 14, no. 1, p. 8254, 2024.

[12] X. Zheng, W. Liu, and Y. Huang, "A novel feature extraction method based on Legendre multi-wavelet transform and auto-encoder for steel surface defect classification," IEEE Access, vol. 12, pp. 5092-5102, 2024.

[13] F. Wang, X. Jiang, Y. Han, and L. Wu, "YOLO-LSDI: An Enhanced Algorithm for Steel Surface Defect Detection Using a YOLOv11 Network," Electronics, vol. 14, no. 13, p. 2576, 2025.

[14] X. Li, C. Xu, J. Li, X. Zhou, and Y. Li, "Multi-scale sensing and multi-dimensional feature enhancement for surface defect detection of hot-rolled steel strip," Nondestructive Testing and Evaluation, vol. 40, no. 8, pp. 3669–3692, 2025.

[15] X. Zheng, W. Liu, and Y. Huang, "Legendre multiwavelet-based feature attention guidance lightweight network for accurate steel surface defect classification," Engineering Applications

of Artificial Intelligence, vol. 161, p. 112179, 2025.
[16] Courtois, J.-M. Morel, and P. Arias, "Investigating Neural Architectures by Synthetic Dataset Design," in Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 4886-4895, 2022.
[17] K. Dikshit, "NEU Surface Defect Database," Kaggle, available: https://www.kaggle.com/datasets/kaustubhdikshit/neu-surface-defect-database/data, accessed Sep. 28, 2025.
[18] K. C. Song and Y. Yan, "A noise-robust method based on completed local binary patterns for hot-rolled steel strip surface defects," Applied Surface Science, vol. 285, pp. 858-864, 2013.
[19] T. TorabiPour, "Foulad Project," GitHub, available: https://github.com/torabi225/foulad/tree/main, accessed Sep. 28, 2025.
[20] T. Torabipour, A. Gandomi, and M. Ghanimi, "A Two-Stage Method for Diagnosing COVID-19, Leveraging CNN, and Transfer Learning on CT Scan Images," International Journal of Web Research, vol. 6, no. 2, pp. 133-142, 2023.
[21] T. TorabiPour, “Steel Surface Defect Detection App,” Streamlit, available: https://foulad77mappmawvcndm87gpzwhca.streamlit.app/, accessed Sep. 28, 2025.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 73 - 88

Steel Surface Defect Classification Using an Improved VGG16-Based Model

Touba Torabipour[1], Abolfazl Gandomi [2*]

[1]PhD Candidate, Department of Computer Engineering, Islamic Azad University, Yazd, Iran

[2]Assistant Professor, Department of Computer Engineering, Islamic Azad University, Yazd, Iran

Article Information

Original Research Paper

Received:
2025 September 05

Accepted:
2025 November 07

Keywords:
Deep Learning, Steel Surface Defect Classification, Improved VGG16 Architecture, Attention Mechanism

Corresponding Author*:
Abolfazl.Gandomi@iau.ac.ir

Abstract

Steel sheets are among the key components used in the automotive industry. Accurate classification of surface defects in these sheets plays a vital role in ensuring product quality. The main challenge in this field lies in the inspection process, as traditional manual inspection methods suffer from significant limitations caused by human error. Although deep learning algorithms have recently emerged as effective and automated approaches for defect detection, the classification of small and complex defects remains challenging and requires optimized modeling of spatial and channel information.In this study, a novel model based on the VGG16 architecture and transfer learning is proposed. The model integrates a two-dimensional spatial attention module and an attention-based magnification mechanism to suppress background noise and focus on the key regions of the image. After evaluation on two benchmark datasets, NEU-CLS and NEU-DET, the proposed model achieved classification accuracies of 99.98% and 99.92%, respectively—showing an improvement of at least 4% over the baseline VGG16 model and approximately 0.018% compared with previous studies. These results demonstrate the superior performance of the proposed approach in accurately classifying steel surface defects. Finally, for practical deployment, a web application was developed and made available to industry experts.

: 10.22034/ABMIR.2025.23563.1162

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۰۵ - ۸۹

مدلی ترکیبی در پیش‌بینی داده محور مصرف انرژی در صنعت فولاد: رویکردی بر پایه LSTM و بهینه‌سازی هایپرپارامترها با الگوریتم Optuna

لیلا ذوالقدر[۱]، مجید شخصی نیائی[۲*]، یحیی زارع مهرجردی[۳]، محمدمهدی لطفی[۳]

[۱] دانشجوی دکتری مهندسی صنایع، گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران

[۲] دانشیار، گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران

[۳] استاد، گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۵/۳۱

تاریخ پذیرش:
۱۴۰۴/۰۷/۲۸

کلیدواژه‌ها:
مدل‌سازی داده محور، LSTM-Optuna ، اعتبار سنجی گام به گام، مصرف انرژی الکتریکی، صنعت فولاد، RNN، SVM

نویسنده‌مسئول:
m.niaei@yazd.ac.ir

چکیده

در دنیای مدرن، انرژی الکتریکی یکی از اساسی‌ترین نیازهای جوامع، نقش حیاتی در بخش‌های صنعتی، اقتصادی و اجتماعی ایفا می‌کند. با توجه به عدم امکان ذخیره‌سازی در مقیاس بزرگ و نوسانات عرضه و تقاضا، پیش‌بینی دقیق مصرف بار الکتریکی ضروری است. مدل‌های سنتی سری‌های زمانی، اغلب در مواجهه با الگوهای غیرخطی و پیچیده داده‌های صنعتی دچار چالش می‌شوند. این پژوهش یک چارچوب ترکیبی نوین برای پیش‌بینی دقیق سری زمانی مصرف انرژی الکتریکی ارائه می‌دهد. رویکرد پیشنهادی بر پایه شبکه عصبی حافظه بلند-کوتاه مدت (LSTM) است که برای به حداکثر رساندن عملکرد مدل، هایپرپارامترهای شبکه LSTM با استفاده از الگوریتم بهینه‌سازی خودکار Optuna بهینه‌سازی شده‌اند. داده‌های مورد استفاده شامل سری‌های زمانی مصرف برق کارخانه فولاد بافت به همراه متغیرهای برونزا مانند میزان تولید فولاد، مدت زمان توقف تولید، شرایط آب و هوایی و تقویمی هستند. نتایج ارزیابی نشان می‌دهند که مدل ترکیبی LSTM-Optuna با ضریب تعیین (R2) ۰/۸۹و درصد خطای مطلق میانگین (MAPE)۱۸/۴۶٪، نه تنها از مدل پایه با ضریب تعیین ۰/۸۵ و درصد خطای مطلق میانگین با مقدار ۲۰/۴۴٪ عملکرد بهتری نشان می‌دهد، بلکه از سایر روش‌های یادگیری ماشین مانند شبکه‌های عصبی بازگشتی (RNN) و ماشین بردار پشتیبان (SVM).

: 10.22034/ABMIR.2025.23569.1157

E-ISSN: 2821-2037

۱- مقدمه

انرژی الکتریکی به‌عنوان یکی از اساسی‌ترین و حیاتی‌ترین نیازهای جوامع بشری، نقشی زیربنایی در توسعه صنعتی، اقتصادی و اجتماعی ایفا می‌کند و امکان ذخیره‌سازی این انرژی در ابعاد بزرگ با فناوری‌های موجود محدود است[۱]. از طرفی انرژی الکتریکی به‌عنوان نیروی محرکه اصلی در صنایع سنگین، نقشی زیربنایی در فرآیندهای تولید ایفا می‌کند. در صنعت فولاد، که فرآیندهای آن به شدت وابسته به کوره‌های قوس الکتریکی و ماشین‌آلات پرمصرف است، مصرف بهینه و پایدار انرژی، ستون فقرات تولید محسوب می‌شود. با توجه به اینکه هزینه‌های انرژی بخش قابل توجهی از هزینه‌های عملیاتی کارخانه‌ها را تشکیل می‌دهد، پیش‌بینی دقیق مصرف برق برای مدیران انرژی و تولید، یک ضرورت راهبردی است.

با این حال، دستیابی به پیش‌بینی دقیق در صنعت فولاد به دلیل وابستگی مصرف به عوامل پیچیده و غیرخطی، یک چالش بزرگ است. الگوهای مصرف انرژی نه تنها تحت تأثیر برنامه‌های تولیدی، بلکه به شدت به متغیرهای برون‌زا و نوسانات ناگهانی و شدید ناشی از عواملی چون قطعی‌های برق برنامه‌ریزی‌نشده و توقف‌های اضطراری نیز وابسته است. این شرایط، داده‌های صنعتی را به مراتب پیچیده‌تر از داده‌های عمومی یا خانگی مصرف برق می‌کند. مدل‌های سنتی سری‌های زمانی، که بر مفروضات خطی بنا شده‌اند، در مواجهه با این الگوهای غیرخطی و نوسانات ناگهانی، کارایی لازم را ندارند [۲]. این امر به‌ویژه در صنایع پرمصرف مانند صنعت فولاد، که مصرف انرژی آن تحت تأثیر متغیرهای برون‌زا قرار می‌گیرد، اهمیت بیشتری پیدا می‌کند[۳]. در سال‌های اخیر، با رشد یادگیری عمیق، رویکردهای نوین مبتنی بر شبکه‌های عصبی عمیق، به‌خصوص مدل حافظه بلند-کوتاه مدت(LSTM) ، به عنوان راهکارهای کارآمد برای پیش‌بینی سری‌های زمانی مطرح شده‌اند[٤].LSTM به دلیل توانایی منحصربه‌فرد خود در یادگیری وابستگی‌های طولانی‌مدت، به‌خوبی می‌تواند الگوهای پیچیده و غیرخطی را در داده‌های مصرف انرژی صنعتی شناسایی کند[٥]. اما چالش فراتر از این است؛ همان‌طور که در پژوهش‌های اخیر بر روی داده‌های صنعتی (مانند حوزه نفت و گاز) نشان داده شده، حتی به‌کارگیری چارچوب‌های پیشرفته یادگیری ماشین خودکار نیز می‌تواند به دلیل پیچیدگی داده‌ها با مشکلات جدی مانند بیش‌برازش مواجه شود و مدل‌هایی غیرقابل اعتماد تولید کند. مطالعه ژو و همکاران (۲۰۲۵) این امر نشان را می‌دهد که صرفاً استفاده از یک مدل پیشرفته کافی نیست و نیاز به یک رویکرد مستحکم و هوشمندانه برای ساخت و اعتبارسنجی مدل در شرایط واقعی صنعتی به شدت احساس می‌شود[٦].

با این حال، عملکرد بهینه این مدل‌ها به تنظیم دقیق هایپرپارامترهای آن‌ها وابسته است [۷]. تنظیم دستی این پارامترها فرآیندی زمان‌بر و غیربهینه است و تضمینی برای دستیابی به بهترین پیکربندی مدل نیز وجود ندارد[۸]. از طرفی با توجه به آن که نوسانات غیرقابل پیش‌بینی و شدید، یک چالش اساسی در مدیریت انرژی و پیش‌بینی دقیق مصرف برق به شمار می‌رود، در این پژوهش، تمرکز بر روی داده‌هایی است که علاوه بر الگوهای فصلی و روزانه، تحت تأثیر عواملی چون قطعی‌های برق برنامه‌ریزی‌نشده قرار دارند. چرا که این امر، دقت مدل‌های پیش‌بینی سنتی را به شدت کاهش می‌دهد و کاملاً تداعی‌کننده شرایط واقعی در صنعت فولاد و محدودیت‌های موجود در مصرف برق است.

این پژوهش در تلاش برای پاسخگویی به تمامی چالش‌های پیش‌بینی مصرف انرژی در صنایع سنگین، یک چارچوب داده‌محور جامع را معرفی می‌کند. این رویکرد پیشنهادی، بر قدرت شبکه عصبی LSTM در مدل‌سازی وابستگی‌های زمانی پیچیده تکیه دارد و آن را با یک الگوریتم بهینه‌سازی هوشمند هایپرپارامترها (Optuna) ترکیب می‌کند. نوآوری و هدف اصلی این تحقیق، نه صرفاً ترکیب دو تکنیک، بلکه اعتبارسنجی و به‌کارگیری این رویکرد پیشرفته بر روی داده‌های واقعی، پیچیده و پرنویز صنعت فولاد است تا نشان دهد آیا چنین مدلی می‌تواند با موفقیت از پس چالش‌های دنیای واقعی صنعت برآید و مدلی دقیق و در عین حال قابل تعمیم ارائه دهد. برای اثبات کارایی این چارچوب، عملکرد آن با مدل پایه و همچنین طیف وسیعی از روش‌های کلاسیک و مدرن پیش‌بینی سری زمانی شامل شبکه‌های عصبی بازگشتی (RNN) و ماشین بردار پشتیبان (SVM) مقایسه شده است.

۲- پیشینه تحقیق

ادبیات پژوهشی در زمینه پیش‌بینی مصرف برق یک مسیر تکاملی طولانی را طی کرده است. در ابتدا، روش‌های آماری و کلاسیک و رگرسیون، ابزارهای اصلی برای این کار بودند. بان و فارمر [۹]و موغرام و رحمان[۱۰]از اولین کسانی بودند که به مرور و مقایسه روش‌های پیش‌بینی کوتاه‌مدت در صنعت برق پرداختند. با این حال، همان‌طور که ژانگ و کی[۲]اشاره کردند، این مدل‌ها در مواجهه با روابط غیرخطی پیچیده با محدودیت روبرو هستند.

با پیشرفت علوم رایانه، مدل‌های مبتنی بر هوش مصنوعی و شبکه‌های عصبی مطرح شدند. این مدل‌ها، به دلیل توانایی در یادگیری الگوهای پیچیده و غیرخطی از داده‌ها، عملکرد بهتری نسبت به روش‌های سنتی نشان داده‌اند. شبکه عصبی مصنوعی (ANN)و پس از آن، مدل‌های پیشرفته‌تری نظیر شبکه‌های عصبی پیچشی (CNN) و شبکه‌های عصبی بازگشتی (RNN) پا به عرصه گذاشتند [۱۱]. مطالعات مختلفی، کارایی شبکه‌های عصبی مصنوعی (ANN) را در پیش‌بینی مصرف برق و سایر متغیرهای صنعتی نشان داده‌اند[۱۲].

در ادامه ماشین بردار پشتیبان (SVM) با معرفی نسخه رگرسیون آن (SVR)، به ابزاری قدرتمند برای پیش‌بینی مقادیر پیوسته تبدیل شد. این مدل به دلیل توانایی در مدیریت فضاهای با ابعاد بالا و مقاومت در برابر بیش‌برازش از طریق تکنیک هسته، در پیش‌بینی سری‌های زمانی، به ویژه در حوزه‌های مالی و انرژی، کاربرد فراوانی دارد [۱۳].

برای حل مشکل محوشدگی گرادیان در RNN ها، هوخرایتر و شمیدهوبر [۱۴]مدل حافظه طولانی-کوتاه‌مدت (LSTM) را معرفی کردند. ژنگ و همکاران[۱۵]و پی و همکاران[۱۶]نیز از LSTM برای پیش‌بینی کوتاه‌مدت با نتایج قابل توجهی استفاده کرده‌اند که به دلیل ساختار حافظه‌دار خود، برای پیش‌بینی سری‌های زمانی بسیار مناسب هستند[۱۷]. تحقیقات متعددی اثبات کرده‌اند که LSTM عملکرد قابل توجهی در پیش‌بینی مصرف انرژی[۱۸]و سایر سری‌های زمانی پیچیده دارد[۱۹،۲۰]. برخی مطالعات به مقایسه LSTM با مدل‌های آماری پرداخته و برتری آن را در شرایط خاص نشان داده‌اند[۲۱].

با وجود کارایی بالای مدل‌های یادگیری عمیق، محققان به این نتیجه رسیده‌اند که یک مدل تکی به‌تنهایی نمی‌تواند تمام پیچیدگی‌های سری‌های زمانی را مدل‌سازی کند. از این رو، رویکردهای ترکیبی توسعه یافتند. یکی از این تکنیک‌ها، تجزیه سیگنال است. هوانگ و همکاران[۲۲]روش تجزیه مد تجربی (EMD) و دراگومیرتسکی و زوسو[۲۳]روش تجزیه مد تغییراتی (VMD) را معرفی کردند. روان و همکاران[۲۴]از یک مدل ترکیبی VMD-LSTM استفاده کردند که نتایج بهتری نسبت به مدل‌های تکی به دست آورد.

همچنین پژوهش‌هایی مانند اکوندایو (۲۰۲۰) و قَرایی و همکاران (۲۰۲۵)، از چارچوب Optuna برای بهینه‌سازی مدل‌های ترکیبی مانند CNN-LSTM در پیش‌بینی مصرف برق استفاده کردند [۲۵،۲۶]. به طور مشابه، یودین و همکاران (۲۰۲۵) با بهینه‌سازی هایپرپارامترهای یک مدل LSTM به خطای بسیار پایینی دست یافتند[۲۷]. هی و همکاران[۷] از بهینه‌سازی بیزی برای بهینه‌سازی LSTMاستفاده کردند که به بهبود قابل توجهی در دقت پیش‌بینی منجر شد.

با نگاهی به پیشینه تحقیق، مشخص می‌شود که تحقیقات به سمت مدل‌های ترکیبی و بهینه‌سازی‌شده حرکت کرده‌اند. مطالعاتی مانند بوکتیف و همکاران [۲۸]از بهینه‌سازی‌های فراابتکاری استفاده کرده‌اند که اغلب از بهینه‌سازی خودکار و هوشمند هایپرپارامترها غفلت کرده‌اند و اکثر تحقیقات مذکور برروی داده‌های عمومی یا خانگی مصرف برق متمرکز شده‌اند و کمتر به چالش‌های منحصربه‌فرد داده‌ها همان‌طور که در پژوهش ژو و همکاران (۲۰۲۵) در حوزه صنعت نفت و گاز نشان داده شد، پرداخته شده است[۶]. چرا که به‌کارگیری مدل‌های یادگیری ماشین پیشرفته برروی داده‌های صنعتی می‌تواند با چالش‌هایی جدی مانند بیش‌برازش همراه باشد.

این پژوهش، با اعتبارسنجی یک چارچوب ترکیبی که به‌طور هم‌زمان از قابلیت‌های یادگیری عمیق LSTM برای مدل‌سازی وابستگی‌های بلندمدت و از یک الگوریتم بهینه‌سازی هوشمند و خودکار مانند Optuna برای تنظیم دقیق هایپرپارامترها در حوزه پیش‌بینی مصرف انرژی بر روی داده‌های پیچیده و پرنویز صنعت فولاد استفاده می‌کند، به دنبال پر کردن این شکاف پژوهشی است.

۳- مفاهیم پایه پژوهش

۳-۱ شبکه‌های عصبی عمیق، مدل حافظه کوتاه‌مدت طولانی LSTM برای سری‌های زمانی

حافظه کوتاه‌مدت طولانی (LSTM) به دلیل توانایی منحصر به فرد خود در یادگیری وابستگی‌های طولانی‌مدت، به یکی از محبوب‌ترین مدل‌ها برای سری‌های زمانی تبدیل شده است [۲۹]. LSTM یک نوع خاص از شبکه عصبی بازگشتی (RNN) است که توسط سپ هوخرایتر و یورگن شمیدهوبر در سال ۱۹۹۷ معرفی شد[۳۰]. یک سلول LSTM، به جای یک تابع فعال‌سازی ساده، دارای ساختار داخلی پیچیده‌تری است که شامل سه دروازه و یک حالت سلول است[۲۹].

۱- حالت سلول: این خط اصلی حافظه سلول است که اطلاعات را در طول توالی منتقل می‌کند.

۲- دروازه فراموشی: این دروازه تصمیم می‌گیرد که چه اطلاعاتی از حالت سلول قبلی (C_{t-1})باید فراموش شوند یا کنار گذاشته شوند. این دروازه یک لایه سیگموئید است که خروجی بین ۰ و ۱ تولید می‌کند. عدد ۱ به معنای نگه‌داشتن کامل و ۰ به معنای کاملاً فراموش کردن است که در آن دروازه فراموشی از رابطه شماره ۱ محاسبه می‌شود:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (۱)$$

که در آن x_tورودی فعلی (مشاهده در زمان t)، h_{t-1} خروجی پنهان از گام زمانی قبلی، W_f ماتریس وزن‌های دروازه فراموشی، b_f بردار بایاس دروازه فراموشی و σ تابع فعال‌سازی سیگموئید است.

۳- دروازه ورودی: این دروازه تصمیم می‌گیرد که چه اطلاعات جدیدی به حالت سلول فعلی (C_t) اضافه شوند. این شامل دو بخش است که یک لایه سیگموئید دیگر که تصمیم می‌گیرد کدام مقادیر آپدیت شوند (دروازهi_t) و یک لایه tanh که یک بردار از کاندیداهای مقادیر جدید $(\tilde{C}_t)$ را ایجاد می‌کند که می‌توانند به حالت سلول اضافه شوند که با توابع نمایش داده شده در رابطه ۲ محاسبه می‌شوند.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (۲)$$
$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

که در آن W_i, W_C ماتریس‌های وزن‌ها، b_i, b_C بردارهای بایاس و tanh تابع فعال‌سازی تانژانت هایپربولیک است.

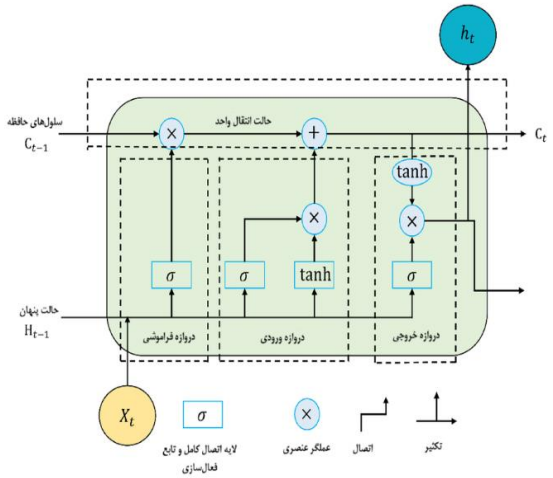
٤- به‌روزرسانی حالت سلول: در این مرحله، حالت سلول قبلی C_{t-1}با ترکیب اطلاعات فراموشی و اطلاعات ورودی جدید، به حالت سلول فعلی C_tبه‌روز می‌شود و در رابطه شماره ۳ ارائه شده است.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (۳)$$

۵- دروازه خروجی: این دروازه تصمیم می‌گیرد که چه بخشی از حالت سلول فعلی C_tبه عنوان خروجی پنهان در این گام زمانی ارائه شود. این خروجی به گام زمانی بعدی منتقل شده و همچنین می‌تواند به عنوان پیش‌بینی مدل استفاده شود. دروازه خروجی o_t و خروجی پنهان h_t با روابط نمایش داده در رابطه ٤ محاسبه می‌شود.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (٤)$$
$$h_t = o_t * \tan h(C_t)$$

که در آن W_o ماتریس وزن‌های دروازه خروجی و b_oبردار بایاس دروازه خروجی است. LSTM ها قادرند وابستگی‌های طولانی‌مدت را یاد بگیرند. این قابلیت کنترلی، به LSTM امکان می‌دهد تا اطلاعات مرتبط را برای مدت‌های طولانی حفظ کند و اطلاعات نامربوط را به سرعت فراموش کند [۲۹،۳۰]. در شکل ۱ معماری داخلی مدل LSTM نمایش داده شده است.



شکل (۱): معماری داخلی مدل حافظه کوتاه‌مدت طولانی LSTM [۳۱]

۳-۲ بهینه‌سازی هایپرپارامترها با استفاده از Optuna

عملکرد مدل‌های یادگیری عمیق، از جمله LSTM، به شدت به تنظیم دقیق هایپرپارامترها وابسته است. هایپرپارامترها، پارامترهایی هستند که قبل از فرآیند آموزش مدل تعیین می‌شوند و بر ساختار و فرآیند یادگیری مدل تأثیر می‌گذارند [۸]. انتخاب نادرست این پارامترها می‌تواند منجر به عملکرد ضعیف مدل، بیش‌برازش یا کم‌برازش شود. هایپرپارامترهای معمول در مدل‌های LSTM شامل مواردی همچون تعداد لایه‌ها و نورون‌ها، نرخ یادگیری[1]، اندازه دسته‌ای[2]، تعداد دوره‌های آموزش[3] و نرخ حذف تصادفی[4] برای جلوگیری از بیش‌برازش، می‌باشند.

روش‌های سنتی برای تنظیم این هایپرپارامترها، مانند جستجوی شبکه‌ای[5] یا جستجوی تصادفی[6]، اغلب ناکارآمد و زمان‌بر هستند، به‌ویژه زمانی که فضای جستجو بسیار بزرگ است. در پاسخ به این چالش، الگوریتم‌های بهینه‌سازی خودکار هایپرپارامترها توسعه یافته‌اند. Optuna یک فریم‌ورک متن‌باز و مدرن برای بهینه‌سازی هایپرپارامترها است که با استفاده از الگوریتم‌های هوشمند و پویا، بهترین ترکیب از پارامترها را پیدا می‌کند [۸،۳۲]. این امر باعث می‌شود که فرآیند بهینه‌سازی سریع‌تر و کارآمدتر انجام شود[۸]. انتخاب Optuna به جای الگوریتم‌های فرابتکاری سنتی (مانند الگوریتم ژنتیک یا ازدحام ذرات) به دلیل کارایی بالاتر آن در فضاهای جستجوی پیچیده است. Optuna از بهینه‌سازی بیزی مبتنی بر مدل جانشین[7] استفاده می‌کند که به آن اجازه می‌دهد تا با هر آزمایش، درک بهتری از تابع هدف پیدا کرده و نواحی امیدوارکننده‌تر را به صورت هوشمندانه کاوش کند. این رویکرد، برخلاف جستجوی عمدتاً تصادفی در روش‌های فرابتکاری، تعداد آزمایش‌های مورد نیاز برای رسیدن به یک راه‌حل بهینه را به شدت کاهش می‌دهد و فرآیند بهینه‌سازی را سریع‌تر و کارآمدتر می‌سازد [۸، ۳۲]. مراحل کار با Optuna به صورت خلاصه به شرح زیر است:

۱- تعریف تابع هدف: تابعی که هایپرپارامترها را به عنوان ورودی دریافت و معیار ارزیابی را به عنوان خروجی بازمی‌گرداند. هدف Optuna کمینه‌سازی یا بیشینه‌سازی این مقدار است.

۲- ایجاد مطالعه: یک شیء Study در Optuna ایجاد می‌شود که فرآیند بهینه‌سازی را مدیریت می‌کند.

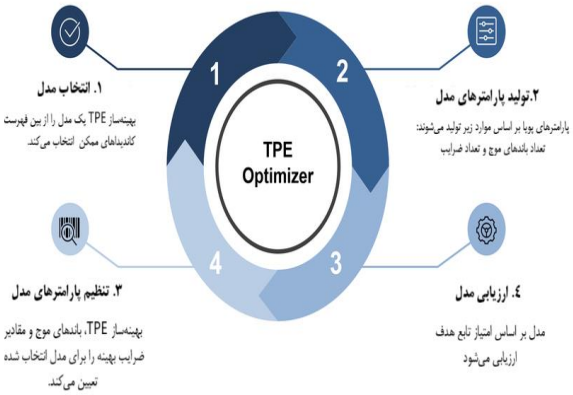
۳- بهینه‌سازی: Optuna به صورت تکراری تابع هدف را با ترکیبات مختلفی از هایپرپارامترها اجرا می‌کند و با استفاده از الگوریتم‌های نمونه‌برداری هوشمند خود، بهترین ترکیب بعدی را پیشنهاد می‌دهد.

۴- تحلیل نتایج: پس از اتمام بهینه‌سازی، Optuna بهترین مجموعه هایپرپارامترها را به همراه بهترین عملکرد مدل ارائه می‌دهد.

Optuna با استفاده از الگوریتم درخت پارزن[7] (TPE) به عنوان مدل جانشین، فضای جستجو را کاوش می‌کند. در هر مرحله، Optuna بهترین نقطه بعدی را برای جستجو با توجه به تعادل بین اکتشاف[8] (جستجوی نواحی جدید و ناشناخته) و بهره‌برداری[9] (تمرکز بر نواحی امیدوارکننده) انتخاب می‌کند. این رویکرد هوشمندانه باعث می‌شود که Optuna نسبت به روش‌های سنتی، سریع‌تر و کارآمدتر عمل کرده و نتایج بهتری را در بهینه‌سازی مدل‌های پیچیده مانند LSTM ارائه دهد.

این الگوریتم، داده‌های تریال‌های قبلی را به دو گروه خوب (دارای عملکرد بالا) و بد (دارای عملکرد پایین) تقسیم کرده و دو مدل چگالی احتمالی[10] را برای هر گروه تخمین می‌زند. این مدل‌ها به ترتیب با $l(x)$ برای گروه خوب و $g(x)$ برای گروه بد نشان داده می‌شوند. در هر مرحله، TPE به دنبال یافتن هایپرپارامترهای جدید x است که تابع اکتساب (E.I.) که در رابطه ۵ ارائه شده است را به حداکثر برسانند. این تابع، تعادل بین اکتشاف و بهره‌برداری را برقرار می‌کند [۳۲].

[1] Learning Rate
[2] Batch Size
[3] Epochs
[4] Dropout
[5] Grid Search
[6] Random Search
[7] Surrogate Model
[8] Exploration
[9] Exploitation
[10] Probability Density Function



شکل (۲): شماتیک عملکرد الگوریتم درخت پارزن (TPE) در بهینه‌سازی هایپرپارامترها با استفاده از Optuna [۳۳]

$$E.I.(x) = \frac{g(x)}{l(x)} \qquad (۵)$$

با استفاده از Optuna ، مدل LSTM قادر خواهد بود به صورت خودکار و بهینه، بهترین پیکربندی را برای پیش‌بینی مصرف انرژی در صنعت فولاد پیدا کند، که این امر به بهبود قابل توجهی در دقت و کارایی مدل منجر خواهد شد [۷۸].

همان‌طور که در شکل ۲ نشان داده شده است، الگوریتم بهینه‌ساز TPE یک فرآیند چرخشی را برای یافتن بهترین هایپرپارامترها دنبال می‌کند. این فرآیند از چهار گام اصلی تشکیل شده است که در گام اول، انتخاب مدل که در آن بهینه‌ساز از میان کاندیداهای ممکن، یک مدل را انتخاب می‌کند. گام دوم تولید پارامترهای مدل که در آن پارامترهای پویا بر اساس توزیع‌های احتمالی قبلی تولید می‌شوند. گام سوم، تنظیم پارامترهای مدل که در آن الگوریتم به صورت هوشمندانه بهترین مقادیر را برای پارامترهای مدل انتخاب می‌کند. در نهایت، ارزیابی مدل انجام می‌شود که در آن مدل بر اساس امتیاز توابع هدف که در اینجا R2 و MAPE است، ارزیابی می‌گردد. این چرخه تا زمانی که بهترین پیکربندی ممکن یافت شود یا تعداد تریال‌ها به پایان برسد، تکرار می‌شود.

۳-۳ معیارهای ارزیابی عملکرد مدل‌ها

۳-۳-۱ ضریب تعیین[1] (R^2)

برای ارزیابی دقت مدل‌های پیش‌بینی، از معیار آماری استاندارد استفاده می‌شود. این معیار میزان برازش مدل به داده‌های واقعی را اندازه‌گیری می‌کنند و امکان مقایسه عادلانه بین مدل‌ها را فراهم می‌کند.R2 یک معیار آماری است که در رابطه شماره ٦ نمایش داده شده است و نسبت واریانس در متغیر وابسته را که از متغیرهای مستقل قابل پیش‌بینی است، نشان می‌دهد. مقدار آن از ۰ تا ۱۰۰٪ (یا ۰ تا ۱) متغیر است، که مقدار بالاتر نشان‌دهنده برازش بهتر مدل به داده‌ها است [۳۳].

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \qquad (٦)$$

که در آن SSresمجموع مربع‌های باقیمانده و SStotمجموع کل مربع‌ها است.

برای ارزیابی دقیق عملکرد مدل‌های پیش‌بینی، از معیارهای ارزیابی کمی استاندارد استفاده می‌شود. این معیارها امکان مقایسه عینی مدل‌ها و سنجش دقت پیش‌بینی آن‌ها را فراهم می‌آورند [۱۲،۳٤].

در این پژوهش، چهار معیار اصلی مورد استفاده قرار گرفته است.

۳-۳-۲ میانگین مربعات خطا[2]

میانگین مربع تفاوت بین مقادیر واقعی Y_iو پیش‌بینی شده $\hat{Y}_i$است که به خطاهای بزرگتر وزن بیشتری می‌دهد و طبق رابطه شماره ۷ محاسبه می‌شود.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad (۷)$$

۳-۳-۳ میانگین خطای مطلق[3]

MAE میانگین تفاوت‌های مطلق بین مقادیر پیش‌بینی شده و واقعی را محاسبه می‌کند و طبق رابطه ۸ محاسبه می‌شود. این معیار خطای پیش‌بینی را در واحدهای داده اصلی ارائه می‌دهد و نسبت به داده‌های پرت کمتر حساس است.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | Yi - \hat{Y}_i | \qquad (۸)$$

۳-۳-٤ ریشه میانگین مربع خطا[4]

RMSE ریشه دوم میانگین مربع خطاهای پیش‌بینی را محاسبه می‌کند. این معیار به خطاهای بزرگ‌تر وزن بیشتری می‌دهد و یک

[1] Coefficient of Determination
[2] Mean Squared Error
[3] Mean Absolute Error
[4] Root Mean Squared Error

دیدگاه جامع از دقت مدل ارائه می‌کند و در رابطه شماره ۹ نمایش داده شده است.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_t - \hat{Y}_t)^2} \quad (۹)$$

۳-۳-۵ میانگین خطای مطلق درصدی[1]

MAPE میانگین خطای مطلق را به صورت درصدی از مقادیر واقعی بیان می‌کند. این معیار امکان مقایسه عملکرد مدل را در میان مجموعه‌های داده مختلف با واحدهای متفاوت فراهم می‌کند و به راحتی قابل تفسیر است و با رابطه شماره ۱۰ محاسبه می‌شود.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \quad (۱۰)$$

۳-٤ اعتبارسنجی مدل

برای ارزیابی عملکرد یک مدل یادگیری ماشین و اطمینان از قابلیت تعمیم‌پذیری آن به داده‌های دیده نشده، از روش‌های اعتبارسنجی متقاطع استفاده می‌شود. این روش‌ها با تقسیم مجموعه داده به زیرمجموعه‌های آموزش و اعتبارسنجی، عملکرد مدل را به صورت پایدار و قابل اعتماد می‌سنجند. رایج‌ترین نوع این تکنیک، اعتبارسنجی متقاطع K-Fold است.

در روش K-Fold، مجموعه داده D به K زیرمجموعه مساوی و تصادفی (F1,F2,...,FK) تقسیم می‌شود، به طوری که هیچ زیرمجموعه‌ای شامل داده‌های قبل یا بعد از زیرمجموعه دیگر نیست. در هر تکرار (k∈{1,...,K})، مجموعه آموزشی (Dtrain,k) و اعتبارسنجی (Dvalid,k) به صورت رابطه‌های ۱۱ و ۱۲ تعریف می‌شود:

$$D_{valid,k} = F_k \quad (۱۱)$$

$$D_{train,k} = D / F_k \quad (۱۲)$$

با این حال، استفاده از K-Fold برای داده‌های سری زمانی کاملاً نامناسب است؛ زیرا ترتیب زمانی داده‌ها را نادیده گرفته و به داده‌های آینده اجازه می‌دهد تا به صورت تصادفی وارد مجموعه آموزش شوند. این پدیده که به عنوان نشت داده[2] شناخته می‌شود، منجر به ارزیابی بیش از حد خوش‌بینانه از عملکرد مدل می‌گردد.

به همین دلیل، در این پژوهش از روش تخصصی اعتبارسنجی متقاطع گام به گام[3] که در کتابخانه‌های پایتون تحت عنوان TimeSeriesSplit پیاده‌سازی شده، استفاده گردید. این روش با حفظ کامل توالی زمانی داده‌ها، یک چارچوب واقع‌گرایانه برای ارزیابی مدل فراهم می‌کند. در روش TimeSeriesSplit، داده‌ها به صورت پیوسته و بدون تداخل زمانی تقسیم می‌شوند. اگر مجموعه داده کامل را به صورت D={x1,x2,...,xN} در نظر بگیریم، در هر تکرار (k∈{1,...,K})، مجموعه آموزشی و اعتبارسنجی به صورت رابطه‌های ۱۳ و ۱٤ تعریف می‌شود:

$$D_{train,k} = \{x_1, x_2, \dots, x_{t_k}\} \quad (۱۳)$$

$$D_{valid,k} = \{x_{t_{k+1}}, x_{t_{k+2}}, \dots, x_{t_{k+N}}\} \quad (۱٤)$$

که در آن tkنقطه پایانی مجموعه آموزشی در تکرار k ام است. این رویکرد تضمین می‌کند که مدل تنها بر روی داده‌های گذشته آموزش دیده و بر روی داده‌های آینده ارزیابی می‌شود. این روش، ارزیابی پایداری مدل را در طول زمان امکان‌پذیر می‌سازد و برای بهینه‌سازی هایپرپارامترها با ابزارهایی مانند Optuna بسیار ضروری است تا از انتخاب پارامترهایی که به صورت تصادفی روی یک برش زمانی خاص عملکرد خوبی داشته‌اند، جلوگیری شود و بهترین مدل به صورت پایدار پیدا شود [۳۵]که این امر اعتبار و قابلیت تعمیم‌پذیری نتایج را در شرایط عملیاتی به‌شدت افزایش می‌دهد.

٤- روش‌شناسی پژوهش

٤-۱ محدوده پژوهش و رویکرد داده‌کاوی

پژوهش حاضر بر پیش‌بینی مصرف انرژی در کارخانه احیای مستقیم فولاد بافت در استان کرمان، شهرستان بافت، تمرکز دارد. این کارخانه با ظرفیت تولید ۸۰۰ هزار تن آهن اسفنجی در سال، از فناوری پرد[4] استفاده می‌کند. این تحقیق از منظر هدف، کاربردی و توسعه‌ای محسوب می‌شود. از منظر زمان جمع‌آوری داده‌ها، پیمایشی است. از بعد ماهیت داده‌ها، کمی بوده و نهایتاً در روش جمع‌آوری داده، از نوع میدانی به شمار می‌رود.

[1] Mean Absolute Percentage Error
[2] Data Leakage
[3] Walk-Forward Cross-Validation
[4] PERED

٤-٢ داده‌ها و منبع آن

پژوهش حاضر بر اساس داده‌های ۵ ساله مصرف انرژی برق کارخانه فولاد بافت صورت گرفته است. این داده‌ها به صورت روزانه از تاریخ ۱ فروردین ۱۳۹۸ تا ۲۹ اسفند ۱۴۰۲ (معادل ۲۱ مارس ۲۰۱۹ تا ۱۹ مارس ۲۰۲٤) در بازه زمانی پنج سال گذشته جمع‌آوری شده‌اند و در مجموع شامل ۱۸۲۷ رکورد می‌باشند. متغیرهای مورد استفاده در این تحلیل شامل مصرف برق به عنوان متغیر هدف و متغیرهای تولید فولاد، ساعات توقف، میانگین، حداقل و حداکثر دما، فشار هوا، رطوبت، تعطیلات رسمی ایران، ویژگی روز هفته به عنوان متغیر برون‌زا در نظر گرفته شده‌اند. داده‌های مربوط به دما، فشار و رطوبت، که به عنوان متغیرهای برون‌زا در مدل‌سازی تأثیرگذار هستند، از سایت‌های جهانی معتبر هواشناسی و اقلیم‌شناسی احصا شده و به مجموعه داده اصلی اضافه گردیده‌اند. این رویکرد داده‌محور، امکان بررسی جامع تأثیر عوامل مختلف بر مصرف انرژی را فراهم می‌کند.

٤-٣ پیش‌پردازش داده‌ها

پیش‌پردازش داده‌ها گامی حیاتی در مدل‌سازی سری‌های زمانی، به ویژه در کاربردهای صنعتی، محسوب می‌شود، چرا که کیفیت داده‌های ورودی تأثیر مستقیمی بر دقت و قابلیت اطمینان مدل‌های پیش‌بینی دارد. در این پژوهش، مراحل پیش‌پردازش داده‌ها به شرح زیر انجام شده است:

۱- تبدیل تاریخ‌ها: ابتدا تاریخ‌های شمسی به میلادی تبدیل و به عنوان ایندکس زمانی دیتافریم تنظیم شده‌اند.

۲- رسیدگی به مقادیر گمشده[1]: مقادیر گمشده در مجموعه داده با استفاده از روش میانگین‌گیری[2] پر شده‌اند.

۳- ایجاد ویژگی‌های جدید[3]: برای مدل‌سازی بهتر تأثیر عوامل زمانی، ویژگی‌های روز هفته (مانند دوشنبه، سه‌شنبه و غیره) با استفاده از روش One-Hot Encoding ایجاد شده‌اند [۱۱].

٤- اعمال تبدیل لگاریتمی و نرمال‌سازی: برای کاهش نوسانات شدید و پایدارسازی سری زمانی، تبدیل لگاریتمی بر روی متغیر هدف (مصرف برق) اعمال شد. سپس، تمامی متغیرها با استفاده از روش نرمال‌سازیMin-Max به مقیاسی بین ۰ و ۱ درآورده شدند تا عملکرد مدل‌های یادگیری عمیق بهبود یابد.

۵- تقسیم داده‌ها[4]: داده‌ها به دو مجموعه آموزش[5]، و آزمون[6] تقسیم شده‌اند و در نهایت ارزیابی نهایی عملکرد مدل بر روی داده‌های دیده نشده (مجموعه آزمون) را فراهم می‌آورد [۱۸].

٤-٤ ابزارهای نرم‌افزاری

پیاده‌سازی این پژوهش با استفاده از زبان برنامه‌نویسی پایتون و ابزارهای متنوعی صورت گرفته است. برای مدیریت داده‌ها و عملیات پیش‌پردازش از کتابخانه Pandas و برای محاسبات عددی از NumPy استفاده شد. مدل‌سازی شبکه‌های عصبی عمیق با استفاده از فریم‌ورک TensorFlow/Keras انجام گردید. همچنین، بهینه‌سازی هایپرپارامترها با استفاده از کتابخانه Optuna و بصری‌سازی نتایج با Matplotlib و Seaborn صورت پذیرفت.

٤-٥ مدل LSTM و تنظیمات آن

در این پژوهش مدل LSTM یک شبکه عصبی متوالی با سه لایه LSTM با ۵۰ نرون[7] در هر لایه است و سه لایه Dropout با نرخ ۰,۲ برای جلوگیری از بیش‌برازش[8] قرار داده شده است. لایه خروجی یک لایه Dense با ۱ واحد برای پیش‌بینی مقدار مصرف انرژی است. این مدل با بهینه‌ساز 'adam' و تابع هزینه 'mean_squared_error'کامپایل شده و برای ۱۵۰ دوره آموزش دیده است. ازEarly Stopping با صبر ۱۰ دوره برای جلوگیری از بیش‌برازش و بازیابی بهترین وزن‌ها استفاده شده است. در جدول ۱ به تشریح تمامی تنظیمات بیان شده است.

[1] Missing Values Imputation
[2] Mean Imputation
[3] Feature Engineering
[4] Data Splitting
[5] Training Set
[6] Test Set
[7] neurons
[8] Overfitting

جدول (۱): تنظیمات مدل شبکه عصبی عمیق LSTM و مدل بهینه Optuna

شاخص	مقدار
بخش داده و مدل‌سازی	
تعداد کل رکوردها	۱۸۲۷
تقسیم داده	۸۰٪ (۱۴۶۱ رکورد) برای آموزش ۲۰٪ (۳۶۶ رکورد) برای آزمون
مدل‌سازی	
نوع مدل	شبکه عصبی LSTM
تابع هزینه	MSE
الگوریتم بهینه‌سازی	Adam
تعداد دوره‌های آموزش	۱۵۰
اندازه دسته	۳۲
بخش بهینه‌سازی(Optuna)	
نوع بهینه‌سازی	بهینه‌سازی دو هدفه
تعداد اهداف	۱. حداکثرسازی ضریب تعیین (R^2) ۲. حداقل‌سازی درصد خطای مطلق میانگین
تعداد تریال‌ها	۵۰
الگوریتم نمونه‌برداری	TPE

ابعاد ورودی این لایه با لایه‌های قبلی متفاوت است. این نشان می‌دهد که لایه آخر LSTM (lstm_14)فقط آخرین خروجی را به لایهDense فرستاده است.

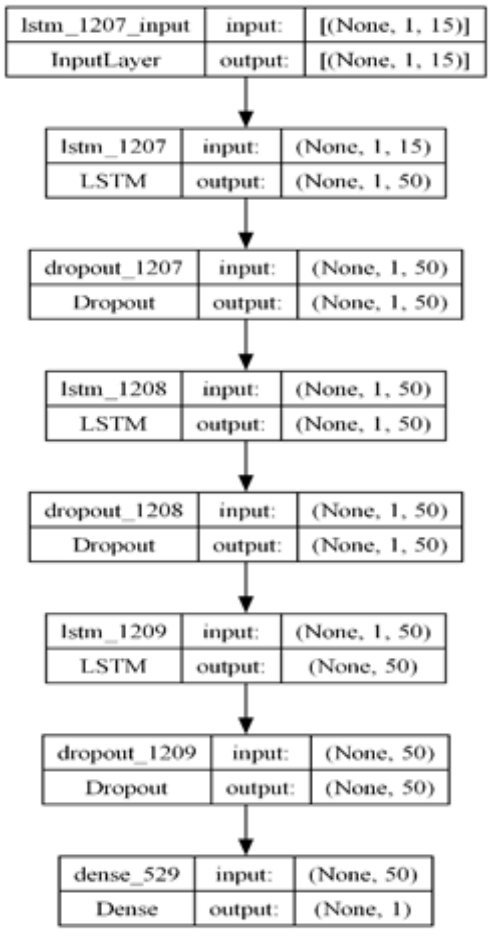
۵- تحلیل نتایج

در این بخش، نتایج حاصل از ارزیابی دو مدل LSTM پایه و LSTM بهینه‌سازی‌شده با Optuna ، براساس معیارهای کمی و بصری تحلیل می‌شوند. مقایسه جامع عملکرد مدل‌ها در دو مرحله پیش از بهینه‌سازی و پس از آن، ارزش رویکرد پیشنهادی را به وضوح نشان می‌دهد.

۵-۱ تحلیل بهینه‌سازی هایپرپارامترها و نمودارهای Loss

مقایسه هایپرپارامترهای به‌دست‌آمده از Optuna با هایپرپارامترهای مدل پایه نشان می‌دهد که بهینه‌ساز تغییراتی استراتژیک و هدفمند را اعمال کرده است. مقادیر مدل پایه و مدل بهینه‌سازی شده در جدول ۲ نمایش داده شده است.

همان‌طور که در جدول مشاهده می‌شود، الگوریتم Optuna با وجود حفظ تعداد لایه‌ها و واحدهای لایه اول، تغییرات کلیدی را در هایپرپارامترهای دیگر ایجاد کرده که به بهبود قابل توجهی در عملکرد مدل منجر شده است. مهم‌ترین تغییری که Optuna ایجاد کرده، افزایش نرخ یادگیری از ۰/۰۰۱ به ۰/۰۰۲۳۷ است. این امر به مدل اجازه داده است تا وزن‌ها را با سرعت بیشتری به‌روز کند و به یک نقطه بهینه در فضای پارامترها دست یابد. تعداد واحدهای لایه سوم را از ۵۰ به ۶۴ افزایش داده و هم‌زمان، نرخ Dropout را در لایه‌های میانی و انتهایی کاهش داده است.



شکل (۳): معماری و اندازه تنسورهای ورودی و خروجی هر لایه شبکه عصبی

این تغییرات نشان می‌دهد که بهینه‌ساز دریافته است که مدل به ظرفیت یادگیری بیشتری در لایه‌های عمیق‌تر نیاز دارد و با کاهش نرخ Dropout، به مدل اجازه داده است تا وابستگی‌های پیچیده‌تر را به طور کامل‌تری یاد بگیرد، بدون اینکه دچار بیش‌برازش شود. این مقایسه به وضوح نشان می‌دهد که بهینه‌سازی خودکار، به جای

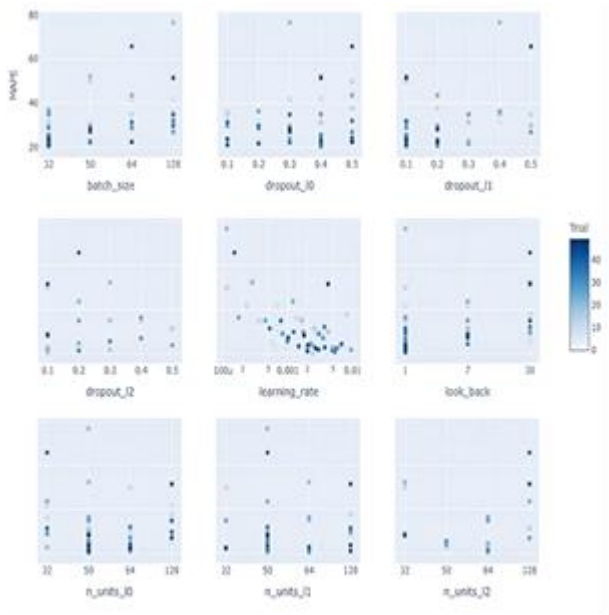
تغییرات کوچک و تصادفی، تغییراتی استراتژیک و هدفمند را در هایپرپارامترها اعمال کرده است. در ادامه نمودار Slice Plot در شکل ٤ به تفکیک هر مورد نمایش داده شده است و تأثیر هر یک از هایپرپارامترها را بر روی معیار درصد خطای مطلق میانگین (MAPE)به صورت جداگانه نمایش می‌دهد.

جدول (۲): مقایسه هایپرپارامترها در مدل شبکه عصبی عمیق LSTM و مدل بهینه Optuna

هایپرپارامتر	مدل پایه	مدل بهینه‌شده باOptuna	تحلیل تغییرات
تعداد لایه‌ها(n_layers)	۳	۳	بدون تغییر
تعداد واحدها در لایه اول (n_units_l0)	۵۰	۵۰	بدون تغییر
نرخ Dropout در لایه اول(dropout_l0)	۰/۲	۰/٤	افزایش یافته
تعداد واحدها در لایه دوم (n_units_l1)	۵۰	۵۰	بدون تغییر
نرخ Dropout در لایه دوم(dropout_l1)	۰/۲	۰/۱	کاهش یافته
تعداد واحدها در لایه سوم (n_units_l2)	۵۰	٦٤	افزایش یافته
نرخ Dropout در لایه سوم(dropout_l2)	۰/۲	۰/۱	کاهش یافته
نرخ یادگیری	۰/۰۰۱	۰/۰۰۲۳۷	افزایش یافته
اندازه دسته	۳۲	۳۲	بدون تغییر

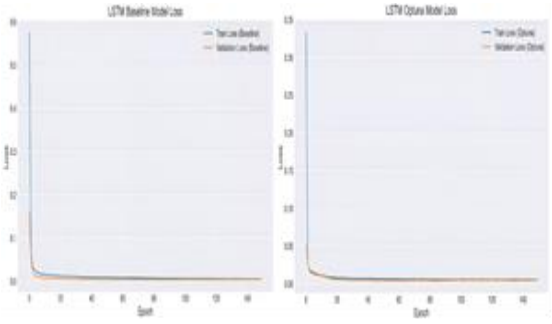
هر نقطه در این نمودار نشان‌دهنده یک تریال بهینه‌سازی است و رنگ نقاط، بیانگر ترتیب زمانی تریال‌ها است. نمودار نرخ یادگیری به وضوح نشان می‌دهد که مقادیر پایین MAPE عمدتاً در یک بازه مشخص از نرخ یادگیری (تقریباً ۰/۰۰۱ تا ۰/۰۰۵) متمرکز شده‌اند. این امر تأیید می‌کند که نرخ یادگیری یکی از حیاتی‌ترین هایپرپارامترها برای دستیابی به دقت بالا در این مدل بوده است. طول پنجره زمانی بهترین نتایج (پایین‌ترین MAPE) به طور مشخص در پنجره زمانی برابر با ۱ به دست آمده‌اند.

این یافته نشان می‌دهد که برای این مجموعه داده خاص، نگاه به یک گام زمانی گذشته، برای پیش‌بینی مقدار بعدی کافی بوده و استفاده از پنجره‌های طولانی‌تر منجر به بهبود عملکرد نشده است. در مجموع، این نمودار به صورت بصری تأیید می‌کند که Optuna به جای جستجوی تصادفی، به صورت هوشمندانه فضای پارامترها را کاوش کرده و توانسته است به ناحیه‌ای از هایپرپارامترها با عملکرد بهینه همگرا شود.



شکل (٤): نمودار Slice Plot و نمایش تأثیر هایپرپارامترها بر معیار درصد خطای مطلق میانگین (MAPE)

در ادامه در بررسی شکل ۵، نمودار روند کاهش خطای مدل (Loss) را در طول فرآیند آموزش برای هر دو مدل نشان می‌دهند.

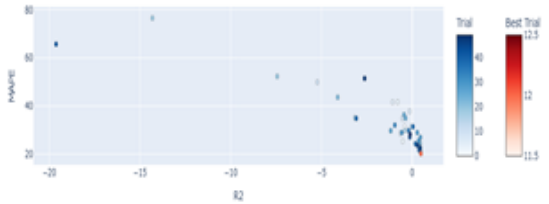


شکل (۵): نمودار تابع هزینه (Loss) مدل پایه در طول فرآیند آموزش

هر دو نمودار همگرایی مناسبی را نشان می‌دهند، اما مدل بهینه‌سازی‌شده از همان ابتدای آموزش، با یک کاهش سریع‌تر و پایدارتر در خطا مواجه شده و به یک سطح خطای بسیار پایین‌تر همگرا می‌شود. این امر نشان می‌دهد که هایپرپارامترهای بهینه شده،

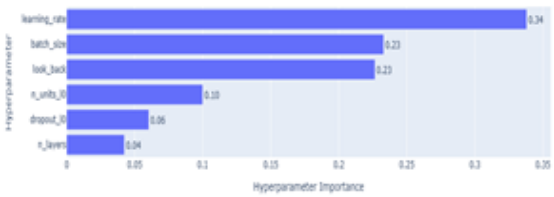
فرآیند یادگیری را کارآمدتر کرده و از نوسانات غیرضروری جلوگیری کرده‌اند.

در تحلیل فرآیند بهینه‌سازی نمودارهای مرز پارتو در شکل ٦، اهمیت هایپرپارامترها و تاریخچه بهینه‌سازی مورد بررسی قرار گرفته است. در نمودار مرز پارتو توازن بین دو هدف بهینه‌سازی (حداکثرسازی R2 و حداقل‌سازی MAPE) را نشان می‌دهد. بهترین تریال‌ها نمایش داده شده که نشان‌دهنده پیدا شدن یک نقطه تعادل بهینه بین دقت و خطای پیش‌بینی است.



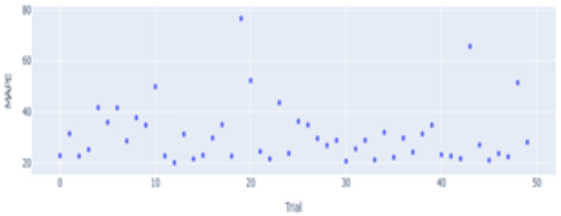
شکل (٦): نمودار مرز پارتو در ایجاد توازن بین دو هدف بهینه‌سازی

نمودار اهمیت هایپرپارامترها[1] تأثیر هر هایپرپارامتر را بر دقت مدل در شکل ۷ نشان می‌دهند. هایپرپارامترهای learning_rate (نرخ یادگیری) و look_back (تعداد گام‌های زمانی گذشته) بیشترین تأثیر را بر روی هر دو معیار R2 و MAPE داشته‌اند. این یافته اهمیت تنظیم دقیق این دو پارامتر را در مدل‌سازی سری‌های زمانی با LSTM برجسته می‌کند.



شکل (۷): نمودار اهمیت هایپرپارامترها و بررسی تأثیر هر هایپرپارامتر بر دقت مدل

نمودار تاریخچه بهینه‌سازی[2] در شکل ۸ روند تغییرات MAPE را در طول ۵۰ تریال بهینه‌سازی نشان می‌دهند. با پیشروی تریال‌ها، الگوریتم Optuna به سمت نواحی با MAPE پایین‌تر حرکت کرده، که نشان‌دهنده کارایی الگوریتم در کاوش هوشمندانه فضای پارامترها است.



شکل (۸): نمودار تاریخچه بهینه‌سازی و بررسی روند تغییرات MAPE در طول ۵۰ تریال

۵-۲ تحلیل عملکرد مدل‌های LSTM و LSTM-Optuna براساس معیارهای کمی

نتایج اعتبارسنجی اولیه مدل پایه با استفاده از روش TimeSeriesSplit، میانگین عملکرد آن را با R2 برابر ۰/٤۸ و MAPEبرابر ۱۹/۵۸٪ نشان داد. این نتایج به عنوان معیاری برای شروع فرآیند بهینه‌سازی در نظر گرفته شدند. پس از آموزش مدل‌ها بر روی کل مجموعه داده آموزش و ارزیابی نهایی در مجموعه آزمون، نتایج به شرح جدول ۳ به دست آمد.

جدول (۳): مقایسه معیارهای ارزیابی در مدل شبکه عصبی عمیق LSTM و مدل بهینه Optuna

معیار ارزیابی	مدل پایه (Baseline)	مدل بهینه‌سازی‌شده با Optuna	بهبود
ضریب تعیین (R^2)	۰/۸۵	۰/۸۹	٪٤/۷۰
درصد خطای مطلق میانگین (MAPE)	٪۲۰/٤٤	٪۱۸/٤٦	٪۹/۷۰
ریشه میانگین مربعات خطا (RMSE)	۹۳۲/۵۵ ,۳۵	۸۲٤/٦۵ ,۳۰	٪۱٤/۱۰
میانگین مربعات خطا (MSE)	۱, ۲۹۱ ,۱٤۷ ۹۵۵/۹۷	,۱۵۸ ,۹۵۰ ۷۵۵/۹۱	٪۲٦/٤۱
میانگین خطای مطلق (MAE)	۸۸۵/۰۲ ,۲٤	۸۵٦/٤۰ ,۲۱	٪۱۲/۱۷

[1] Hyperparameter Importances

[2] Optimization History

همان‌طور که در جدول ۳ مشاهده می‌شود، مدل LSTM-Optuna در تمامی معیارهای ارزیابی به طور قابل توجهی بهتر از مدل پایه عمل کرده است. افزایش ضریب تعیین از ۰/۸۵ به ۰/۸۹ و کاهش ۱/۹۸ درصدی در MAPE نشان‌دهنده توانایی تبیینی بالاتر و دقت بیشتر مدل بهینه‌سازی‌شده است.

در نهایت با مقایسه بصری پیش‌بینی‌ها در شکل ۹، مقایسه بصری عملکرد دو مدل را بر روی مجموعه آزمون نهایی نشان می‌دهد. همان‌طور که در نمودار مشاهده می‌شود، پیش‌بینی‌های مدل بهینه‌سازی‌شده با Optuna (خط قرمز) به طور چشمگیری انطباق نزدیک‌تری با مقادیر واقعی (خط آبی) دارند. به خصوص در دوره‌هایی که نوسانات شدید یا افت ناگهانی در مصرف انرژی وجود دارد، مدل بهینه‌سازی‌شده دقت بالاتری در دنبال کردن الگوهای واقعی نشان می‌دهد. این مقایسه بصری، برتری مدل ترکیبی را در مواجهه با داده‌های پیچیده و غیرخطی صنعتی تأیید می‌کند.

شکل (۹): مقایسه بصری پیش‌بینی مدل شبکه عصبی عمیق LSTM و مدل بهینه Optuna

۵-۳ مقایسه عملکرد با مدل‌های دیگر

برای ارزیابی جامع و اثبات برتری مدل پیشنهادی، عملکرد آن با چندین مدل شناخته‌شده دیگر در حوزه پیش‌بینی سری زمانی مقایسه گردید. این مدل‌ها شامل شبکه عصبی بازگشتی ساده، ماشین بردار پشتیبان ساده و بهینه‌شده بودند. تمامی مدل‌ها بر روی یک مجموعه داده یکسان آموزش دیده و ارزیابی شدند. نتایج کمی این مقایسه در جدول ۵ ارائه شده است.

جدول (٤): مقایسه عملکرد مدل پیشنهادی با سایر مدل‌های پیش‌بینی

معیار ارزیابی	RNN	SVR	SVR بهینه شده	LSTM	LSTM-Optuna
ضریب تعیین (R^2)	۰/۸	۰/٤٤	۰/۸۵	۰/۸۵	۰/۸۹
درصد خطای مطلق میانگین (MAPE)	٪۲۱/۰۸	٤۳/۰۲	٪۲٤/٤۹	٪۲۰/٤٤	٪۱۷/٦٤
ریشه میانگین مربعات خطا (RMSE)	۸٦٦/٤٤/۳۳	٥۷۰/۲۹۰/۲	۲۱٤/۹۸/۳۸	۹۳۲/٥٥/۳٥	۸۲٤/٦٥/٤۷

نتایج جدول ٤ به وضوح نشان می‌دهد که مدل پیشنهادی LSTM-Optuna در تمامی معیارهای کلیدی عملکرد بهتری نسبت به سایر مدل‌ها داشته است. مدل RNN به دلیل ساختار حافظه‌دار خود عملکردی منطقی از خود نشان داد، اما نتوانست با نسخه پیشرفته‌تر خود یعنی LSTM رقابت کند. مدل SVR بهینه شده نیز اگرچه توانست بخشی از الگوها را یاد بگیرد، اما در مواجهه با نوسانات شدید و وابستگی‌های زمانی پیچیده داده‌های صنعتی، با محدودیت مواجه شد. این مقایسه جامع، ارزش ترکیب معماری قدرتمند LSTM با یک فرآیند بهینه‌سازی هوشمند هایپرپارامترها را برای دستیابی به بالاترین سطح از دقت در پیش‌بینی‌های صنعتی تأیید می‌کند.

٦- نتیجه‌گیری

پژوهش حاضر، با موفقیت یک چارچوب ترکیبی داده‌محور بر پایه LSTM و بهینه‌سازی چندهفته هدفه با Optuna را برای پیش‌بینی دقیق مصرف انرژی در صنعت فولاد معرفی و ارزیابی کرد. نتایج به وضوح برتری قابل توجه این مدل را نسبت به مدل پایه نشان داد و تأکید کرد که تنها استفاده از شبکه‌های عصبی عمیق کافی نیست و بهینه‌سازی هایپرپارامترها، عاملی حیاتی در رسیدن به عملکرد بهینه است.

با توجه به نتایج حاصل، در حالی که محققانی مانند ژنگ و همکاران [۱۵]و پی و همکاران[۱٦]ارزش مدل‌های LSTM را در

پیش‌بینی سری‌های زمانی اثبات کردند، پژوهش حاضر نشان داد که ترکیب این مدل با یک رویکرد بهینه‌سازی هوشمند، دقت را به سطحی جدید ارتقا می‌دهد و به R2 برابر ۰/۸۹ و MAPE با مقدار ۱۸/۴۶٪ می‌رسد. نتایج ارزیابی نشان داد که مدل بهینه‌سازی‌شده LSTM-Optuna ، از مدل پایه LSTM و همچنین مدل‌های دیگری مانند RNN، SVR و SVR بهینه شده عملکرد بهتری دارد. پژوهش حاضر، با تأیید یافته‌های هی و همکاران[۷]و گائو و همکاران[۸]در خصوص اهمیت بهینه‌سازی هایپرپارامترها، این رویکرد را با استفاده از فریم‌ورک مدرن Optuna در یک کاربرد صنعتی خاص به کار گرفت. مقایسه هایپرپارامترهای بهینه‌شده با مدل پایه نشان داد که Optuna به جای تغییرات کوچک، تغییراتی استراتژیک در نرخ یادگیری و ظرفیت مدل (تعداد واحدها و نرخ Dropout) ایجاد کرده که منجر به همگرایی سریع‌تر و دقت بالاتر شده است. این تحقیق نشان داد که گنجاندن متغیرهای برون‌زای صنعتی و اقلیمی، به همراه داده‌های مربوط به قطعی‌های برنامه‌ریزی‌نشده، برای مدل‌سازی دقیق نوسانات شدید مصرف انرژی در صنعت فولاد ضروری است. این امر تأییدکننده یافته‌های پاشایی و دهخوارقانی[۳]در مورد اهمیت عوامل مؤثر بر مصرف انرژی در صنایع پرمصرف است.

تحلیل مقایسه‌ای در چشم‌انداز پژوهش‌های اخیر برای ارزیابی دقیق‌تر جایگاه این دستاورد، نتایج پژوهش حاضر با یافته‌های جدیدترین مقالات در حوزه پیش‌بینی مصرف انرژی در جدول شماره ۵ مقایسه گردید. پژوهش‌های اخیر مانند یودین و همکاران (۲۰۲۵) و قرایی و همکاران (۲۰۲۵)، با به‌کارگیری مدل‌های LSTM بهینه‌سازی‌شده بر روی داده‌های استاندارد مصرف برق، به دقت‌های بسیار بالایی (به ترتیب R^2=۰/۹۴ و R^2=۰/۹۹) دست یافته‌اند [۲۷،۲۶]. هرچند در نگاه اول، این ضرایب تبیین قدرت کامل مدل پیش‌بینی را به نمایش می‌گذارد، اما تفاوت ماهوی در پیچیدگی و ماهیت مجموعه داده‌ها، نقطه کلیدی تحلیل است. مقالات مذکور بر روی داده‌های عمومی یا خانگی کار کرده‌اند که الگوهای فصلی و روزانه مشخصی دارند. در مقابل، قدرت اصلی مدل حاضر در عملکرد موفق بر روی یک مجموعه داده واقعی، صنعتی و بسیار پرچالش نهفته است. داده‌های مورد استفاده در این پژوهش، شامل نوسانات شدید ناشی از فرآیندهای تولید و به‌ویژه قطعی‌های برق برنامه‌ریزی‌نشده بود که مدل‌سازی را به مراتب دشوارتر می‌کند.

جدول (۵): مقایسه تحلیلی نتایج پژوهش با مطالعات برجسته اخیر

پژوهش	مدل اصلی	حوزه داده	R^2 (آزمون)	MAPE (آزمون)	نکات کلیدی تحلیل
پژوهش حاضر	LSTM + Optuna (دو هدفه)	صنعت فولاد (با قطعی برق)	۰/۸۹	٪۱۸/۴۶	عملکرد قوی و پایدار بر روی داده‌های صنعتی بسیار پرنویز
Uddin et al. (2025)	LSTM + HPO	مصرف برق عمومی	۰/۹۴	٪۲/۶۳	دقت بالا بر روی داده‌های استاندارد و کم‌نویز
Gharai et al. (2025)	LSTM + Optuna	مصرف برق خانگی	۰/۹۹	گزارش نشده	دقت بسیار بالا بر روی داده‌های استاندارد و قابل پیش‌بینی
Zhu et al. (2025)	AutoML	صنعت نفت و گاز	۰/۲۲	گزارش نشده	نمونه‌ای از چالش بیش‌برازش در داده صنعتی

اهمیت این موضوع زمانی برجسته‌تر می‌شود که نتایج پژوهش ژو و همکاران (۲۰۲۵) را در نظر بگیریم. ایشان با استفاده از یک چارچوب پیشرفته یادگیری ماشین خودکار بر روی داده‌های صنعتی نفت و گاز، با مشکل جدی بیش‌برازش مواجه شدند، به طوری که R^2 مدل از ۰/۷۹ بر روی داده‌های آموزش به ۰/۲۳ بر روی داده‌های آزمون سقوط کرد. در مقابل، مدل پیشنهادی ما با دستیابی به R^2 پایدار ۰/۸۹ بر روی داده‌های آزمون، نشان‌دهنده استحکام۱ و قدرت تعمیم‌پذیری۲ بالای آن در شرایط واقعی و

[1] Robustness

[2] Generalization

پرنویز صنعتی است [٦]. علاوه بر این، رویکرد بهینه‌سازی در این پژوهش با در نظر گرفتن هم‌زمان دو هدف (کاهش MAPE و افزایش R^2)، به دنبال یافتن یک مدل متوازن و جامع بود؛ رویکردی که مشابه آن در پژوهش بیجو و پیلای (۲۰۲۳) برای یافتن توازن بهینه بین «دقت» و «تفسیرپذیری» به کار رفته و نشان‌دهنده همسویی روش‌شناسی ما با جدیدترین روندهای علمی است[۳٦].

در نهایت، این پژوهش با ارائه یک مدل دقیق و قابل اعتماد، به مدیران انرژی در صنعت فولاد کمک می‌کند تا یک استراتژی مدیریت انرژی فعال اتخاذ کنند. پیش‌بینی‌های دقیق، امکان برنامه‌ریزی بهینه تولید، مذاکره هوشمندانه‌تر در خصوص قراردادهای تأمین انرژی و جلوگیری از جریمه‌های ناشی از تجاوز از پیک مصرف را فراهم می‌آورد. همچنین، این مدل به مدیران کمک می‌کند تا تأثیر عوامل مختلف بر مصرف انرژی را بهتر درک کرده و برای حفظ پایداری عملیاتی در برابر نوسانات بازار و شرایط محیطی، آمادگی داشته باشند.

علیرغم دستاوردهای مهم این پژوهش، محدودیت‌هایی وجود دارد که می‌تواند به عنوان فرصت‌هایی ارزشمند برای تحقیقات آینده در نظر گرفته شود:

۱- مدل‌های ترکیبی پیشرفته‌تر: پژوهش‌های آتی می‌توانند رویکردی فراتر از ترکیب ساده LSTM با بهینه‌سازی را دنبال کنند. به عنوان مثال، کاوش در مدل‌های ترکیبی مبتنی بر تجزیه سیگنال مانند VMD-LSTM یا ترکیب مدل‌های Encoder-Decoder با مکانیزم توجه[1] می‌تواند به بهبود دقت پیش‌بینی، به ویژه در الگوهای پیچیده‌تر، منجر شود.

۲- بهینه‌سازی پویا و خودکار: در این مطالعه، از Optuna برای بهینه‌سازی پارامترهای اصلی مدل استفاده شد. تحقیقات آینده می‌توانند بهینه‌سازی‌های پیشرفته‌تری را بررسی کنند که به صورت خودکار معماری مدل را نیز تعیین می‌کنند یا از الگوریتم‌های بهینه‌سازی دیگری مانند Hyperopt برای مقایسه عملکرد استفاده نمایند.

۳- افزایش تفسیرپذیری: مدل‌های یادگیری عمیق اغلب به عنوان "جعبه سیاه "شناخته می‌شوند. برای افزایش اعتماد مدیران به نتایج، پژوهش‌های آتی می‌توانند از روش‌های تفسیرپذیری مدل‌های هوش مصنوعی (XAI) مانند SHAP یا LIME استفاده کنند تا دلایل پشت هر پیش‌بینی، به صورت شفاف مشخص شود.

٤- پیش‌بینی چند گام جلوتر و عدم قطعیت: این مطالعه بر پیش‌بینی کوتاه‌مدت متمرکز بود. توسعه چارچوب‌های مدل‌سازی که به طور هم‌زمان پیش‌بینی چند گام جلوتر را با دقت بالا و همراه با تخمین عدم قطعیت ارائه می‌دهند، می‌تواند کاربردهای عملیاتی گسترده‌تری داشته باشد.

سپاسگزاری

این پژوهش با حمایت مالی ارزشمند دانشگاه یزد و شرکت توسعه فناوری و نوآوری فولاد مبارکه اصفهان به سرانجام رسید. نویسندگان از همکاری و پشتیبانی این نهادها در قالب گرنت پارسا در اجرای پروژه پایان‌نامه دکتری با محوریت حل یک چالش کلیدی صنعتی، صمیمانه سپاسگزاری می‌کنند.

References

[1] S. Ghasaee and R. Ravanmehr, “Short-Term Prediction of Electrical Load Consumption Using Deep Neural Networks, CNN, and LSTM,” J. Qual. Product. Iran Electr. Ind., vol. 10, no. 1, pp. 35–51, 2021 [In Persian].

[2] G. Zhang and M. Qi, “Neural Network Forecasting for Seasonal and Trend Time Series,” Eur. J. Oper. Res., vol. 160, no. 2, pp. 501-514, 2005, doi: 10.1016/j.ejor.2003.08.037.

[3] H. Pashaei and M. Dehkharghani, “Predictive Modeling of Energy Consumption in the Steel Industry Using CatBoost Regression: A Data-Driven Approach for Sustainable Energy Management,” J. Clean. Prod., vol. 401, p. 136706, 2023, doi: 10.1016/j.jclepro.2023.136706.

[4] A.Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network,” Physica D: Nonlinear Phenomena, vol. 404, p. 132306, 2020, doi: 10.1016/j.physd.2020.132306.

[5] S. Arslan, “A Hybrid Forecasting Model Using LSTM and Prophet for Energy Consumption with

[1] Attention

Decomposition of Time Series Data," PeerJ Comput. Sci., vol. 8, p. e1001, 2022, doi: 10.7717/peerj-cs.1001.
[6] R. Zhu, N. Li, Y. Duan, G. Li, G. Liu, F. Qu, et al., "Well-Production Forecasting Using Machine Learning with Feature Selection and Automatic Hyperparameter Optimization," Energies, vol. 18, no. 1, p. 99, dic. 2024, doi: 10.3390/en18010099.
[7] F. He, J. Zhou, Z.-k. Feng, G. Liu, and Y. Yang, "A Hybrid Short-Term Load Forecasting Model Based on Variational Mode Decomposition and Long Short-Term Memory Networks, Considering Relevant Factors with a Bayesian Optimization Algorithm," Appl. Energy, vol. 237, pp. 108–117, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.055.
[8] T. Gao, Y. Sun, and C. Li, "LSTM Network Hyperparameter Optimization for Stock Price Prediction Using the Optuna Framework," J. Phys.: Conf. Ser., vol. 1785, no. 1, p. 012061, 2021, doi: 10.1088/1742-6596/1785/1/012061.
[9] D. Bunn and E. Farmer, "Review of Short-Term Forecasting Methods in the Electric Power Industry," in Comparative Models for Electrical Load Forecasting, D. Bunn and E. Farmer, Eds. Berlin: Springer, 1985, pp. 13-30.
[10] Moghram and S. Rahman, "Analysis and Evaluation of Five Short-Term Load Forecasting Techniques," IEEE Trans. Power Syst., vol. 4, no. 4, pp. 1484-1491, Nov. 1989.
[11] K. Amasyali and N. M. El-Gohary, "A Review of Data-Driven Building Energy Consumption Prediction Studies," Renew. Sustain. Energy Rev., vol. 81, pp. 1192-1205, Jan. 2018.
[12] M. H. L. Lee, Y. C. Ser, G. Selvachandran, P. H. Thong, L. Cuong, L. H. Son, et al., "A Comparative Study of Forecasting Electricity Consumption Using Machine Learning Models," Mathematics, vol. 10, no. 1329, pp. 1-23, 2022.
[13] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer, 2017.
[14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
[15] Y. Zheng, J. Zhang, and Y. Wang, "Short-Term Load Forecasting Based on a Long Short-Term Memory Neural Network," J. Mod. Power Syst. Clean Energy, vol. 5, no. 2, pp. 296-302, Mar. 2017.
[16] S. Pei, H. Qin, L. Yao, Y. Liu, C. Wang, and J. Zhou, "Multi-Step Ahead Short-Term Load Forecasting Using Hybrid Feature Selection and Improved Long Short-Term Memory Network," Energies, vol. 13, no. 16, p. 4121, Aug. 2020, doi: 10.3390/en13164121.
[17] M. Zohaki Raht and H. Sadeghi Saghadol, "Modeling and Forecasting of Iran's Short-Term Electricity Consumption Using Neural Network and TPE Algorithm," Journal of Energy Economics Studies, vol. 20, no. 83, pp. 159-182, 2024 [In Persian].
[18] J. Zhu, Y. Wang, C. Lai, and X. Zhou, "Deep Learning-Based Energy Consumption Prediction Model for Green Industrial Parks," Appl. Artif. Intell., vol. 39, no. 1, p. 2462375, 2021.
[19] L. Shen, W. Chen, and J. Kwok, "Deep Learning for Short-Term Energy Consumption Forecasting of Smart Buildings Considering Microgrid and External Factors," Mathematics, vol. 10, no. 1329, pp. 1-23, 2022.
[20] S. Sayadinejad, A. Esmailzadeh Maghari, and M. R. Rostami, "Presenting a Bitcoin Return Prediction Model Using a Hybrid Deep Learning Signal Decomposition Algorithm (CEEMD-DL)," Quarterly Journal of Financial Economics, vol. 17, no. 62, pp. 217-238, 2022 [In Persian].
[21] J. Kim, H. Kim, H. Kim, D. Lee, and S. Yoon, "A Comprehensive Survey of Deep Learning for Time Series Forecasting: Architectural Diversity and Open Challenges," Artif. Intell. Rev., vol. 58, p. 216, Apr. 2025.
[22] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, et al., "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis," Proc. R. Soc. Lond. A, vol. 454, no. 1971, pp. 903-995, Mar. 1998.
[23] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," IEEE Trans. Signal Process., vol. 62, no. 3, pp. 531-544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.
[24] Y. Ruan, G. Wang, H. Meng, and F. Qian, "A Hybrid Model for Power Consumption Forecasting Using VMD-Based Long Short-Term Memory Neural Network," Front. Energy Res., vol. 9, p. 772508, Feb. 2022, doi: 10.3389/fenrg.2021.772508.
[25] I. O. Ekundayo, "Optuna optimization-based CNN-LSTM model for predicting electric power energy consumption," M.S. thesis, School of Comput., Nat. College of Ireland, Dublin, 2020.
[26] C. Gharai, S. K. Mohapatra, S. Parida, C. Dora, R. K. Mohanta, and S. Chakravarty, "Optimized Deep Learning Model for Power Consumption Prediction Using Hyperparameter Tuning Techniques," in Int. Conf. Intell. Cloud Comput.

(ICoICC), Bhubaneswar, India, 2025, pp. 1-6, doi: 10.1109/ICoICC64033.2025.11052111.
[27] R. Uddin, M. R. Hazari, S. Ahmad, C. A. Hossain, M. S. Rahman Zishan, and A. Ahmed, "Short-Term Load Forecasting Using Deep Learning Algorithms with Hyperparameter Optimization," in 4th Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST), Dhaka, Bangladesh, 2025, pp. 366-371, doi: 10.1109/ICREST63960.2025.10914446.
[28] S. Bouktif, R. F'Haim, and M. Lazaar, "A Novel Hybrid Model for Short-Term Load Forecasting Using LSTM and Metaheuristic Optimization Algorithms," Energy, vol. 196, p. 117042, Apr. 2020.
[29] J. Brownlee, Deep Learning for Time Series Forecasting: Theory to Practice. San Juan, PR: Machine Learning Mastery, 2018.
[30] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA: MIT Press, 2016.
[31] J. Zhao, G. Nie, and Y. Wen, "Monthly Precipitation Prediction in Luoyang City Based on EEMD-LSTM-ARIMA Model," Water Sci. Technol., vol. 87, no. 1, pp. 318–335, Jan. 2023, doi: 10.2166/wst.2022.425.
[32] Y. Ren, Z. Yan, D. Liu, J. Hou, and W. Zhou, "Optimized EWT-Seq2Seq-LSTM with Attention Mechanism for Insulator Fault Prediction," Energies, vol. 15, no. 2, p. 528, Jan. 2022, doi: 10.3390/en15020528.
[33] A. Alghamdi and A. Desuqi, "Predicting Electricity Consumption Using Machine Learning Techniques," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 11, pp. 384-391, 2020.
[34] M. Alimohammadi Ardakani, M. H. Karimi-Zarchi, and D. Shishebori, "A Hybrid Forecasting Model for Accurate Prediction of Building Energy Consumption Based on Multi-Stage Decomposition and Optimized Deep Learning," Energy, vol. 305, p. 126955, 2024.
[35] R. J. Hyndman and G. Athanasopoulos. (2021). Forecasting: Principles and Practice (3rd ed.) [Online]. Available: https://otexts.com/fpp3/
[36] B. G. M. and G. N. Pillai, "Hyperparameter Optimization of Long Short Term Memory Models for Interpretable Electrical Fault Classification," IEEE Access, vol. 11, pp. 123688-123704, Nov. 2023, doi: 10.1109/ACCESS.2023.3330056.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 89 - 105

A Hybrid Model for Data-Driven Energy Consumption Forecasting in the Steel Industry: An Approach Based on LSTM and Hyperparameter Optimization with the Optuna Algorithm

Leila Zolqadr[1], Majid Shakhsi-Niaei [2*], Yahia Zare Mehrjerdid [3], Mohammad Mehdi Lotfi[3]

[1]Ph.D. Student in Industrial Engineering, Department of Industrial Engineering, Yazd University, Yazd, Iran

[2]Associate Professor, Department of Industrial Engineering, Yazd University, Yazd, Iran

[3]Professor, Department of Industrial Engineering, Yazd University, Yazd, Iran

Article Information

Original Research Paper

Received:
2025 August 22

Accepted:
2025 October 20

Keywords:
Data-Driven Modeling, STM-Optuna, Walk-Forward Validation, Electrical Energy Consumption, Steel Industry, RNN, SVM

Corresponding Author*:

m.niaei@yazd.ac.ir

Abstract

In the modern world, electrical energy, as one of the most fundamental needs of societies, plays a vital role in the industrial, economic, and social sectors. Given the impossibility of large-scale storage and fluctuations in supply and demand, accurate forecasting of electrical load consumption is essential. Traditional time-series models often struggle to handle the nonlinear and complex patterns of industrial data. This research presents a novel hybrid framework for accurately forecasting electrical energy consumption time series. The proposed approach is based on a Long Short-Term Memory (LSTM) neural network; to maximize the model's performance, the LSTM network's hyperparameters have been optimized using the Optuna automatic optimization algorithm. The data used includes time series of electricity consumption from the Baft steel factory, along with exogenous variables such as steel production volume, production downtime, and weather and calendar conditions. The evaluation results indicate that the hybrid LSTM-Optuna model, with a coefficient of determination (R^2) of 0.89 and a Mean Absolute Percentage Error (MAPE) of 18.46%, not only demonstrates better performance than the baseline model (with an R^2 of 0.85 and a MAPE of 20.44%) but also outperforms other machine learning methods such as Recurrent Neural Networks (RNN) and both simple and optimized Support Vector Machines (SVM).

: 10.22034/ABMIR.2025.23569.1157

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۲۳-۱۰۷

تشخیص خودکار شمش‌های گرد ناصاف از محصولات خروجی دستگاه صافکاری با استفاده از روش‌های بینایی ماشین

راستین نمیرانیان[۱]، سحرسلطانی[۲]، ولی درهمی[۳*]

[۱] واحد فناوری اطلاعات، شرکت فولاد آلیاژی ایران، یزد، ایران

[۲] کارشناسی ارشد هوش مصنوعی و رباتیکز، شرکت توسعه فولاد آلیاژی ایرانیان، یزد، ایران

[۳] استاد، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۷/۰۳

تاریخ پذیرش:
۱۴۰۴/۰۸/۲۳

کلیدواژه‌ها:
بینایی ماشین، تبدیل هاف دایره‌ای، دستگاه صافکاری، شمش فولادی

نویسنده‌مسئول:
vderhami@yazd.ac.ir

چکیده

این مقاله به مساله تشخیص شمش فولادی گرد ناصاف در حال غلتیدن می پردازد. در سالن تکمیل کاری کارخانه فولاد آلیاژی ایران یکی از عملیاتی که روی شمش های گرد انجام می‌شود، صاف کردن آن‌ها است. لیکن هیچ سیستمی برای اینکه کیفیت خروجی دستگاه صافکاری را مشخص کند وجود ندارد. در اینجا برای حل این چالش ویدیوها از شمش خارج شده از دستگاه صافکاری که در حال غلتیدن روی میز است، گرفته می‌شود. پس از پیش پردازش و تعادل کردن نوری تصویر، مقطع شمش در هر تصویر با روش انتقال هاف دایره مشخص می‌شود. سپس مسیر حرکت شمش در هر ویدیو تعیین می‌شود. شش ویژگی از این مسیر حرکت برای مطالعه انتخاب شده است. این ویژگی‌ها شامل : طول مسیر حرکتی، مجموع مربعات انحراف، مجموع مربعات انحراف عمودی، ارتعاشِ انحراف معیار، انرژی ارتعاش مبتنی بر موجک و انحراف معیار شتاب هستند. مقدار این ویژگی ها با لحاظ کردن نمونه برداری در واحد طول محاسبه می‌شوند. آزمایش‌ها نشان می‌دهد که ویژگی‌هایی همچون انحراف معیار ارتعاش، مجموع مربعات انحراف عمودی و مجموع مربعات انحراف، قدرت تمایز بالایی در تفکیک شمش‌های سالم و ناصاف دارند. هرچند مقدار منفی نادرست کمی بالا است ولی روش پیشنهادی توانسته با دقت حدود ۸۵ درصد، عملکرد قابل قبولی ارائه دهد.

: 10.22034/ABMIR.2025.23720.1171

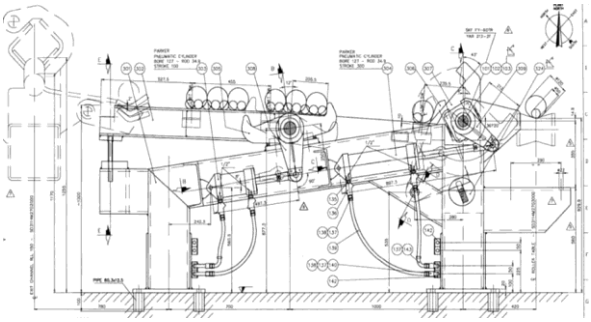
E-ISSN: 2821-2037

۱- مقدمه

فولاد یکی از کالاهای اساسی و استراتژیک در صنایع هر کشور به شمار می‌رود. میزان تولید و مصرف فولاد نه تنها یک شاخص مهم در توسعه صنعتی محسوب می‌شود، بلکه به‌عنوان معیاری برای سنجش رشد اقتصادی و زیرساختی کشورها نیز در نظر گرفته می‌شود. ویژگی‌های ممتاز فولاد از جمله استحکام بالا، شکل‌پذیری مناسب، مقاومت در برابر سایش و قابلیت بازیافت، موجب شده است که این ماده در حوزه‌های متعددی مانند ساخت‌وساز، صنایع انرژی، خودروسازی، تولید تجهیزات صنعتی و حتی صنایع نظامی جایگاه ویژه‌ای داشته باشد. از این رو، کیفیت محصولات فولادی به‌طور مستقیم بر کارایی و ایمنی پروژه‌های صنعتی و عمرانی اثرگذار بوده و هر گونه نقص در فرآیند تولید می‌تواند پیامدهای جدی به دنبال داشته باشد[۱، ۲].

در این مقاله، مساله بررسی میزان صاف بودن شمش‌های فولادی گرد در قسمت تکمیل کاری کارخانه فولاد آلیاژی ایران مورد بحث قرار گرفته است. در کارخانه فولاد آلیاژی ایران، پس از انجام عملیات نورد گرم یا عملیات حرارتی روی شاخه‌های فولادی با سطح مقطع گرد، شمش‌ها وارد مرحله‌ای حیاتی به نام صافکاری دوغلتکی می‌شوند. این فرآیند به‌منظور اصلاح کجی یا انحنای احتمالی شمش‌ها طراحی شده است تا محصول نهایی با بالاترین کیفیت ممکن به بازار عرضه شود. کیفیت اجرای این مرحله تأثیر مستقیمی بر سطح کیفی نهایی دارد و هرگونه خطا در صافکاری، می‌تواند موجب کاهش بهره‌وری و افزایش ضایعات شود.

دستگاه صافکاری دوغلتکی که در شکل ۱ نمایش داده شده است، از سه بخش اصلی تشکیل شده است: میز ورودی، مجموعه صافکاری و میز خروجی. در این مجموعه، شمش‌های فولادی بین دو تیغه صاف و دو غلتک محدب و مقعر از بالا و پایین قرار می‌گیرند. تنظیم زاویه و فشار این غلتک‌ها بسته به سختی شمش و میزان کجی ورودی انجام می‌شود. طراحی خاص غلتک‌ها موجب می‌شود که شاخه علاوه بر حرکت طولی، دوران ملایمی داشته باشد تا با اصلاح کرنش‌های الاستیک، کجی موجود برطرف گردد. در صورتی که فشار کافی اعمال نشود، بخشی از کجی در خروجی باقی می‌ماند و در مقابل، فشار بیش از حد نیز سبب افزایش مصرف انرژی و استهلاک دستگاه خواهد شد.



شکل (۱): شماتیک دستگاه صافکاری دوغلتکی

در حال حاضر در کارخانه، تشخیص صحت این فرآیند و اندازه‌گیری میزان کجی باقیمانده در شمش خروجی، به طور کامل به قضاوت چشمی اپراتورهای انسانی متکی است. این نیروهای خبره با مشاهده‌ی نحوه‌ی غلتیدن شمش بر روی میز خروجی شیب‌دار، سالم یا ناصاف بودن آن را تشخیص می‌دهند. این روش ذاتاً ذهنی و مستعد خطای انسانی است که می‌تواند به خروج محصول معیوب یا کاهش بهره‌وری منجر شود. ضمن اینکه فعلا حضور نیروی خبره انسانی بطور دائم در محل میسیر نیست.

لذا مساله تشخیص میزان ناصافی با روش‌های هوشمند در کارخانه مطرح و مورد پژوهش قرار گرفت. در مسیر تحقیق در مراحل اولیه، ایده‌های مختلفی همچون ثبت و تحلیل صدای شمش در هنگام غلتیدن یا نصب لودسل [۳] برای سنجش نیروهای اعمال‌شده بر میز خروجی مطرح شد. اما این روش‌ها یا به تجهیزات اضافی نیاز داشتند و یا در برابر شرایط محیطی حساس بودند. در نهایت، تیم تحقیقاتی به سمت بهره‌گیری از روش‌های بینایی ماشین هدایت شد، چرا که این روش‌ها امکان ثبت مستقیم حرکت سطح مقطع شمش و استخراج ویژگی‌های مرتبط با میزان صافی یا ناصافی را فراهم می‌کردند.

بر اساس آخرین دانش ما، این مقاله اولین تحقیق در زمینه‌ی تشخیص ناصافی شمش‌ها با استفاده از تحلیل حرکت آن‌ها بر روی میز دشارژ درمحیط باز سالن کارخانه با روش بینایی ماشین است. ایده اصلی بر این فرض استوار است که شمش در خروجی دستگاه، حداقل بخشی از مسیر شیب‌دار را می‌غلتد و امکان ردیابی حرکت سطح مقطع آن فراهم می‌شود. همان‌گونه که در شکل ۲ مشاهده

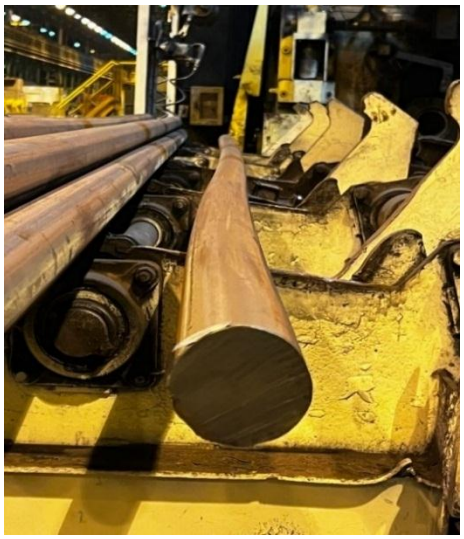
می‌شود، زمانی که میز خروجی پر از شمش‌های دیگر است، فضای کافی برای غلتیدن وجود ندارد و تحلیل حرکت قابل انجام نیست. در مقابل، شکل ۳ شرایطی را نشان می‌دهد که میز خروجی تقریباً خالی است و شمش می‌تواند آزادانه حرکت کند؛ این حالت مبنای اصلی تحلیل‌های پژوهش حاضر قرار گرفته است.

شکل (۲): نمایی از میز خروجی کاملا پر

شکل (۳): نمایی از میز خروجی تقریبا خالی

از طریق ردیابی مسیر حرکت سطح مقطع شمش در چنین شرایطی می‌توان اطلاعات دقیقی در خصوص میزان صافی یا ناصافی به‌دست آورد. ویژگی‌های استخراج‌شده از این مسیر، شامل شاخص‌های حرکتی و ارتعاشی، معیارهای عینی برای تفکیک شمش سالم از ناصاف ارائه می‌دهند. این رویکرد، جایگزینی مطمئن برای قضاوت ذهنی اپراتورهای انسانی است و امکان تصمیم‌گیری علمی، تکرارپذیر و خودکار را در فرآیند کنترل کیفیت فراهم می‌کند.

با وجود جذابیت و کارایی این ایده، چندین چالش اساسی در پیاده‌سازی آن وجود دارد. نخست اینکه میزان ناصافی در برخی شمش‌ها بسیار جزئی است، به‌طوری‌که حرکت آن‌ها شباهت زیادی به شمش‌های سالم دارد و استخراج ویژگی‌های متمایزکننده کار دشواری می‌شود. به‌منظور درک بهتر سطوح مختلف ناصافی، سه نمونه از شمش‌های گرد با درجات متفاوت انحراف مورد بررسی قرار گرفته‌اند. در شکل ٤ نمونه‌ای از شمش خیلی ناصاف و در شکل ۵ نمونه‌ای از شمش کمی ناصاف نمایش داده شده است. همان‌طور که در تصاویر مشاهده می‌شود، در نمونه‌ی خیلی ناصاف، خمیدگی و تاب‌خوردگی در طول شمش به‌وضوح قابل تشخیص است، در حالی که در شمش کمی ناصاف، میزان انحراف بسیار کم بوده و حتی توسط افراد خبره نیز به‌سختی قابل شناسایی است.



شکل (٤): شمش خیلی ناصاف با خمیدگی و تاب‌خوردگی مشخص در طول شمش

شکل (۵): شمش کمی ناصاف، میزان انحراف بسیار کم

دوم اینکه رفتار شمش‌ها بر روی میز خروجی همیشه قابل پیش‌بینی نیست؛ برخی شمش‌ها به دلیل ناصافی یا شرایط سطح میز، در

مسیر کج حرکت می‌کنند یا حتی متوقف می‌شوند. این رفتار غیرخطی می‌تواند داده‌های جمع‌آوری‌شده را نویزی و غیرقابل اتکا سازد.

از سوی دیگر، شرایط تصویربرداری در محیط کارخانه نیز چالش بزرگی به شمار می‌آید. نورپردازی ضعیف باعث شد ویدیوهای گرفته شده اولیه کیفیت مناسبی نداشته باشند و نیاز به انجام پردازش‌های پیشرفته مانند افزایش کنتراست و بهبود وضوح باشد. همچنین، تهیه یک مجموعه داده بزرگ و دقیق که شامل برچسب‌های سالم و ناصاف باشد، نیازمند صرف زمان و هزینه قابل توجه است، با توجه به نبود ابزار دقیق مرتبط برای اندازه‌گیری ناصافی و همچنین محدودیت‌های عملی موجود در کارخانه، فرآیند برچسب‌گذاری نمونه‌ها صرفاً بر پایه قضاوت دو فرد خبره انجام شد.

در این مقاله روش تشخیص ناصافی شمش گرد با فرض غلتیدن در یک حداقل طول با استخراج اطلاعات و ویژگی های مسیر حرکت وسط سطح مقطع شمش ارائه می‌شود.

ساختار این مقاله بدین شرح است: ابتدا در بخش دوم مروری بر کارهای گذشته و در بخش سوم مبانی نظری مرتبط با تشخیص حرکت و تحلیل ارتعاشات ارائه می‌شود. در بخش چهارم، فرآیند گام‌به‌گام روش پیشنهادی شامل پردازش ویدیوها، تشخیص سطح مقطع شمش‌ها، استخراج ویژگی‌های حرکتی و ارتعاشی و نهایتاً روش‌های تصمیم‌گیری تشریح می‌گردد. در بخش پنجم ، ارزیابی نتایج شامل تحلیل آستانه‌گذاری، نمودار مشخصه عملکرد سیستم[1] و روش‌های خوشه‌بندی[2] مانند K-Means ارائه و مقایسه می‌شود. در پایان، جمع‌بندی دستاوردهای پژوهش همراه با پیشنهادهایی برای تحقیقات آینده در بخش ششم مطرح خواهد شد.

۲- مروری بر کارهای گذشته

در این بخش به برخی از کارهای مرتبط با موضوع تحقیق پرداخته می‌شود . چن و همکارانش [٤] روی تشخیص صافی میلگرد تراش خورده سیلندرها به روش سه نقطه ای در حوزه زمان پرداخته‌اند.

در این روش، برای بالابردن دقت از یک سیستم اندازه‌گیری تماسی استفاده شده و مدل‌سازی ریاضی آن استخراج شده است. در نهایت به دقت اندازه‌گیری به کمتر از ۰،۰۰۰۵ میلی‌متر رسیده است. نتایج تحلیل این سیستم نشان می‌دهد که تکرارپذیری خوبی دارد. این روش نیاز به بستر خاص دارد که برای دستگاه صافکاری تحقیق مذکور جوابگو نمی‌باشد.

در تحقیق دیگر سیف و همکارانش [۲] به اندازه گیری گردی محصول می‌پردازند و از تکنیک‌های بینایی ماشین و پردازش تصویر برای تحلیل دقیق هندسه قطعات استفاده کرده‌اند. این روش از الگوریتم‌های پردازش تصویر، مانند تبدیل هاف[3] و تشخیص لبه [4] بهره می‌برد تا ویژگی‌های دایره‌ای قطعات را شناسایی و تحلیل کند. سیستم طراحی شده با استفاده از دوربین‌های با وضوح بالا و نورپردازی مناسب توانسته اختلاف اندازه‌گیری بین ۵ تا ۹/٦ میکرومتر نسبت به سیستم سنتی ماشین اندازه گیری مختصات داشته باشد، که نشان‌دهنده دقت قابل قبول آن در کاربردهای صنعتی است.

در تحقیق دیگر، چن و همکارانش [۱] یک چارچوب جامع برای تشخیص خودکار انحنا و ناصافی در شاخه‌های U شکل ارائه دادند. این پژوهش با ترکیب هوشمندانه‌ای از تکنیک‌های پردازش تصویر و بینایی ماشین، رویکردی چندمرحله‌ای را طراحی کرده است. در مرحله اول، از الگوریتم یولو[5] برای تشخیص دقیق موقعیت شاخه‌ها و استخراج نقاط مرکز به عنوان مرجع استفاده شده است. سپس با به کارگیری روش‌های تشخیص لبه و خط، کانتورهای دقیق استخراج شده و خطوط مربوطه شناسایی می‌شوند. در ادامه، با استفاده از الگوریتم‌های خوشه‌بندی، خطوط هم‌راستا گروه‌بندی شده و با تکنیک دوخت خطوط[6]، قطعات خطی به صورت پیوسته ترکیب می‌شوند. در نهایت، با تحلیل هندسی این خطوط و اندازه‌گیری انحنا، وضعیت تختی یا انحنای شاخه‌ها با دقت بالا تشخیص داده می‌شود.

برای اندازه گیری صافی شمش، روش‌های دیگری مثل کار اشمید و تیمشان [۵] نیز انجام شده است که با استفاده از شدت نور

[1] Receiver Operating Characteristics (ROC curve)
[2] Clustering
[3] Hough Transform
[4] Edge Detection
[5] YOLO (You Only Look Once)
[6] Stitching

بازتاب‌شده نسبت به زاویه اسکن، میزان انحراف از حالت کاملاً مستقیم را محاسبه می‌کند. این تکنیک نیاز به چرخش کامل میله ندارد و فقط در مرحله کالیبراسیون، چرخش جزئی لازم است. مزیت اصلی این روش، امکان استفاده در خطوط تولید بدون تماس فیزیکی با قطعه است. برای انجام کالیبراسیون مورد نیاز در این روش، ممکن است لازم باشد میله یکبار حول محور خود بچرخد تا ویژگی‌های سطح و موقعیت نسبی لیزر مشخص شود. مهمترین چالش در این روش فراهم‌سازی محیط آزمایش و نورپردازی است. چرا که انعکاس نورها تاثیر زیادی روی کیفیت عملکرد روش دارد. در تحقیق دیگر، به مسئله‌ی حیاتی تشخیص خودکار عیوب سطح فولاد پرداخته شده است که در صنایع حساس مانند هوافضا، خودروسازی و دفاعی نقش اساسی در کنترل کیفیت دارد. هدف آن‌ها رفع مشکل تعادل بین دقت بالا و سرعت پردازش واقعی است؛ چالشی که روش‌های بازرسی دستی از عهده‌ی آن برنمی‌آیند. برای این منظور، مدلی بهبودیافته از یولو پنج پیشنهاد شد که توانایی شناسایی شش نوع عیب سطحی فولاد را دارد. این روش با بهره‌گیری از چند نوآوری فنی مانند الگوریتم K-means++ برای بهینه‌سازی کادرهای محصورکننده، تابع خطای اشتراک بر اتحاد کارآمد[1] برای بهبود همگرایی، روش نمونه‌برداری افزایشی کاراف[2] برای حفظ اطلاعات معنایی در گسترش نقشه‌های ویژگی، و مکانیزم توجه[3] برای افزایش حساسیت کانال‌ها، عملکرد مدل پایه را به‌طور چشمگیری ارتقا داده است. این پژوهش نشان می‌دهد که ترکیب روش‌های هوشمند در طراحی شبکه‌های عمیق می‌تواند بدون افزایش زیاد هزینه‌ی محاسباتی، کیفیت تشخیص را در کاربردهای واقعی صنعتی به‌طور مؤثری بهبود بخشد[٦].

در تحقیق دیگر، به مشکل اندازه‌گیری میزان تختی در ریل‌های راه‌آهن پرداخته شده است که برای رفع محدودیت‌های روش‌های سنتی که نیازمند چندین دوربین یا لیزر و حساس به تغییرات نور و بازتاب سطح بودند، یک سامانه اندازه‌گیری خودکار مبتنی بر بینایی ماشین و دوسامانه‌ی مثلث‌سازی لیزری پیشنهاد کردند. این روش با استفاده از کالیبراسیون هم‌زمان برای دو واحد تصویربرداری، خطاهای هندسی را کاهش داده و نگهداری سیستم را ساده‌تر می‌کند. همچنین با به‌کارگیری الگوریتم‌های ثبت پروفایل مقاوم در سه مرحله (تراز اولیه، تراز کلی و تراز محلی)، اثر لرزش‌ها و جابه‌جایی‌های حین تولید جبران می‌شود. در مرحله‌ی تحلیل، از روش‌های محاسبه انحراف سطح و فیلترینگ هوشمند برای جداسازی نویز و استخراج ویژگی‌های واقعی تختی استفاده شده است[٧].

همچنانکه دیده شد هرچند تحقیقاتی روی شمش‌های فولادی انجام شده است، اما تحقیقی بابت تشخیص صافی شمش گرد در محیط باز سالن کارخانه با روش بینایی ماشین انجام نشده است. البته می توان از ایده‌های مطرح شده در مقاله ها بهره برد.

۳- مفاهیم پایه

با توجه به دانش فرد خبره، نحوه غلتیدن و الگوی ارتعاش شمش در هنگام حرکت بر روی میز خروجی می‌تواند شاخصی مهم برای تشخیص میزان صافی یا ناصافی سطح محصول دستگاه صافکاری باشد. نخستین گام در طراحی سامانه تشخیص خودکار، شناسایی دقیق سطح مقطع شمش در حال حرکت و ردیابی مرکز آن در طول ویدیو است. برای این منظور، روش‌های مختلفی نظیر یولو، تبدیل هاف دایره‌ای[4] و تطبیق الگو[5] بررسی شدند و مزایا و محدودیت‌های هرکدام مورد ارزیابی قرار گرفت در این بخش مختصری از این روش‌ها گفته می‌شود.

۳-۱ الگوریتم یولو

الگوریتم یولو یکی از معماری‌های پیشرو در زمینه تشخیص اشیاء مبتنی بر یادگیری عمیق است. این روش نخستین‌بار در سال ۲۰۱۵ توسط جوزف ردمون و همکارانش معرفی شد و با کنار گذاشتن رویکردهای دو مرحله‌ای سنتی مانند شبکه‌های عصبی کانولوشنی مبتنی بر ناحیه[6] ، تحولی بزرگ در سرعت و دقت تشخیص ایجاد کرد [٨]. برخلاف مدل‌هایی که ابتدا نواحی پیشنهادی را شناسایی و سپس آن‌ها را طبقه‌بندی می‌کردند، یولو کل فرآیند را در یک گام و به‌صورت انتها به انتها انجام می‌دهد[٩].

[1] Efficient Intersection over Union (EIoU)
[2] Carafe Upsampling
[3] SE-Net (Squeeze-and-Excitation Network)
[4] Hough Circle Transform
[5] Template Matching
[6] Region-based Convolutional Neural Network

پیشرفت‌های پیاپی در نسخه‌های مختلف یولو نشان‌دهنده تلاش برای بهبود توازن بین سرعت و دقت است. در نسخه‌های جدیدتر، شبکه‌های عمیق‌تری به‌عنوان ستون فقرات معماری استفاده شده‌اند تا استخراج ویژگی‌های دقیق‌تر انجام شود. علاوه بر این، سازوکارهایی نظیر کادرهای محصورکننده برای بهبود تشخیص اشیاء با اندازه‌های مختلف و مکانیزم‌های بهینه‌سازی به‌منظور کاهش نرخ خطا به کار گرفته شده است. این ویژگی‌ها موجب شد یولو گزینه‌ای جذاب برای سامانه‌های نظارت تصویری، خودروهای خودران و بازرسی‌های صنعتی باشد [۹، ۱۰].

با وجود مزایای فراوان، یولو محدودیت‌هایی نیز دارد. به عنوان مثال، تشخیص اشیاء بسیار کوچک یا اشیائی که به شدت همپوشانی دارند ممکن است با دقت کافی انجام نشود. علاوه بر این، برای آموزش مؤثر این مدل به مجموعه داده‌های حجیم نیاز است و در صورتی که کلاس‌های جدید اضافه شوند، باید شبکه از ابتدا یا حداقل به‌صورت جزئی بازآموزی گردد. با این حال، معماری انعطاف‌پذیر و سرعت بالای آن سبب شده است که همچنان یکی از پرکاربردترین ابزارها در حوزه بینایی ماشین باشد.

۳-۲ تبدیل هاف دایره‌ای

تبدیل هاف یکی از ابزارهای کلاسیک و قدرتمند در بینایی ماشین برای شناسایی اشکال هندسی ساده نظیر خطوط و دایره‌ها است. نسخه ویژه این الگوریتم که برای تشخیص دایره طراحی شده است که ایده اصلی آن مبتنی بر نگاشت نقاط موجود در تصویر به فضای پارامتری شعاع و مرکز دایره است. هر نقطه از لبه‌های تصویر به مجموعه‌ای از دایره‌های ممکن نگاشت می‌شود و نقاطی که در فضای پارامتری بیشترین تراکم رأی را دارند به عنوان مراکز احتمالی دایره شناسایی می‌شوند[۱۱].

این الگوریتم بدون نیاز به داده‌های آموزشی حجیم و صرفاً با تحلیل ویژگی‌های هندسی تصویر کار می‌کند، بنابراین برای کاربردهای صنعتی که داده‌های برچسب‌خورده کمی وجود دارد بسیار مناسب است. همچنین به دلیل ماهیت ریاضی روشن، می‌توان پارامترهای آن نظیر آستانه‌های تشخیص لبه یا بازه تغییرات شعاع را به‌طور مستقیم تنظیم کرد. نقطه قوت اصلی تبدیل هاف دایره‌ای در شرایطی است که شکل مقطع شمش‌ها به‌طور طبیعی نزدیک به دایره باشد و محیط تصویربرداری از وضوح مناسبی برخوردار باشد.

با وجود مزیت‌های ذکر شده ، تبدیل هاف دایره‌ای در برابر نویز و تغییرات روشنایی حساس است. به همین دلیل، پیش‌پردازش‌هایی مانند متعادل‌سازی هیستوگرام[1] برای بهبود کنتراست یا اعمال فیلترهای حذف نویز[2] قبل از اجرای الگوریتم اهمیت ویژه‌ای دارد. در این پژوهش، به دلیل سرعت بالا، عدم نیاز به مدل‌سازی داده و قابلیت پیاده‌سازی ساده روی خطوط تولید واقعی، از تبدیل هاف دایره‌ای به‌عنوان گزینه اصلی تشخیص سطح مقطع استفاده شد.

۳-۳ تطبیق الگو

تطبیق الگو یکی از روش‌های سنتی در پردازش تصویر است که بر مبنای یافتن بیشترین شباهت بین یک الگو و بخش‌های مختلف تصویر عمل می‌کند. این الگوریتم با محاسبه معیارهایی مانند همبستگی[3] یا مجموع قدرمطلق اختلاف‌ها میزان شباهت هر ناحیه از تصویر را با الگوی ورودی اندازه‌گیری می‌کند[۱۲]. در صورتی که سطح مقطع شمش همیشه دارای اندازه و روشنایی نسبتاً ثابت باشد، این روش می‌تواند در تشخیص موقعیت مرکز سطح مقطع کارآمد باشد.

از مزیت‌های الگوریتم تطبیق الگو، سادگی و عدم نیاز به مدل‌سازی پیچیده است. تنها کافی است یک الگوی مرجع از سطح مقطع شمش در اختیار داشته باشیم تا با حرکت پنجره‌ای کوچک روی تصویر و محاسبه میزان تطابق، موقعیت دایره پیدا شود. اما این روش در برابر تغییرات اندازه [4] یا چرخش[5] بسیار آسیب‌پذیر است و به همین دلیل در شرایط واقعی خطوط تولید که شمش‌ها ممکن است با سرعت‌های مختلف حرکت کنند یا تحت تأثیر نورهای متفاوت قرار گیرند، دقت و پایداری بالایی ندارد.

یکی دیگر از محدودیت‌های این روش سرعت نسبتاً پایین آن در پردازش ویدئوهای با نرخ فریم بالا است، زیرا باید شباهت در هر فریم به صورت کامل محاسبه شود. همین موضوع باعث شده است

[1] Contrast-Limited Adaptive Histogram Equalization
[2] Gaussian Blur
[3] Correlation
[4] Scaling
[5] Rotation

که در پژوهش حاضر، الگوریتم تطبیق الگو صرفاً به عنوان یک گزینه مطالعاتی در نظر گرفته شود اما به دلیل ضعف عملکرد در محیط پرنویز و تغییرپذیر کارخانه، در نسخه نهایی کنار گذاشته شد.

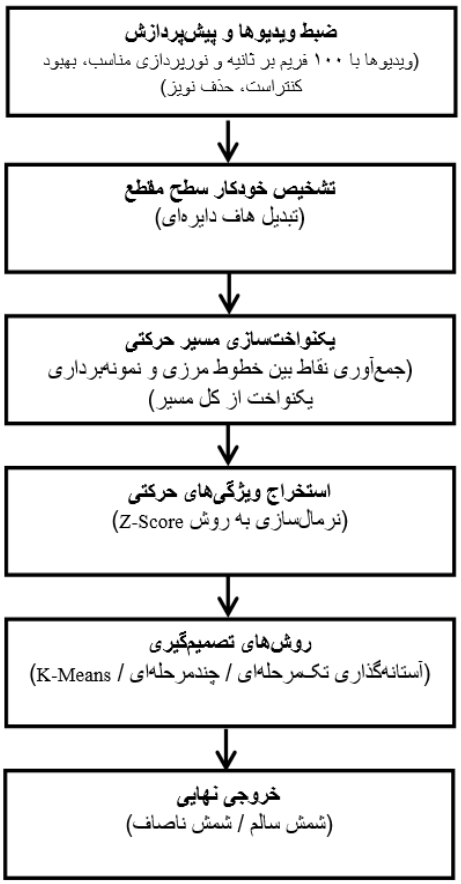
٤- روش پیشنهادی

در این پژوهش روشی برای ارزیابی سلامت شمش‌های فولادی هنگام حرکت روی میز خروجی ارائه شد. ایده اصلی این روش مبتنی بر ثبت و تحلیل مسیر حرکت سطح مقطع شمش در ویدیوهای صنعتی است. مشاهده شد که شمش‌های سالم رفتار دینامیکی یکنواخت‌تری داشته و تغییرات ناگهانی کمتری در مسیر حرکت نشان می‌دهند، در حالی که شمش‌های ناصاف یا دارای ناهمواری، بیشتر دچار انحراف و ارتعاش می‌شوند. این تفاوت‌ها به کمک استخراج ویژگی‌های حرکتی کمی‌سازی شده و برای تصمیم‌گیری خودکار به کار رفت.

فرآیند پیشنهادی به‌گونه‌ای طراحی شده که علاوه بر کارخانه مورد مطالعه، در خطوط تولید مشابه و حتی صنایع دیگر که به پایش حرکت اجسام نیاز دارند، قابل استفاده باشد. به‌طور کلی این روش شامل جمع‌آوری ویدیوها و پیش‌پردازش، تشخیص خودکار سطح مقطع، یکنواخت‌سازی نقاط مسیر، استخراج ویژگی‌های حرکتی و در آخر روش‌های تصمیم‌گیری است. نمای کلی روش پیشنهادی در شکل ٦ نشان داده شده است. در ادامه، هر یک از مراحل به تفصیل در زیر بخش‌های ٤-١ تا ٤-٥ توضیح داده شده‌اند.

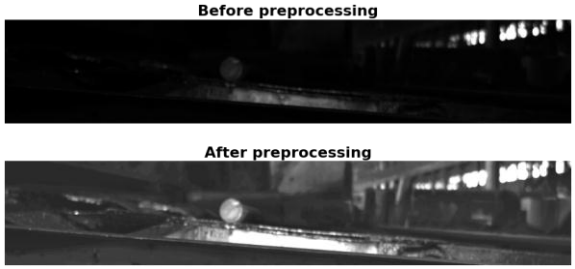
٤-١ جمع‌آوری ویدیوها و پیش‌پردازش

ویدیوهای صنعتی با نرخ فریم بالا برابر با ۱۰۰ فریم بر ثانیه و نورپردازی مناسب جمع‌آوری شدند. همچنین از سرعت بالای شاتر جهت جلوگیری از تاری تصویر و افزایش وضوح حرکت سطح مقطع استفاده شد. برای تأمین روشنایی یکنواخت، دو منبع نورLED با توان ۲۰ وات در پشت دوربین و در دو سوی آن نصب شدند تا نور از زاویه‌ی مایل بر سطح شمش بتابد. سپس ویدیوهای ثبت‌شده تحت پردازش‌های ابتدایی قرار گرفتند.



شکل (٦): نمای کلی روش پیشنهادی برای تشخیص ناصافی شمش گرد

در این مرحله به‌منظور بهبود کیفیت تصویر و حذف اثرات تغییرات نور محیط، روش‌هایی مانند افزایش کنتراست و تنظیم روشنایی اعمال گردید. این پیش‌پردازش‌ها موجب شد تشخیص لبه‌ها و سطح مقطع شمش‌ها پایدارتر و دقیق‌تر انجام گیرد و تغییرات ناخواسته نور محیط کارخانه کاهش یابد. در شکل ۷ نمونه‌ای از اعمال پیش‌پردازش‌های تصویر، نمایش داده شده است. تصویر سمت بالا مربوط به محیط با نور کارخانه است و تصویر سمت پایین تغییرات پس از اعمال پیش‌پردازش‌ها می‌باشد.



شکل (۷): نتایج حاصل از پیش‌پردازش

٤-٢ تشخیص خودکار سطح مقطع

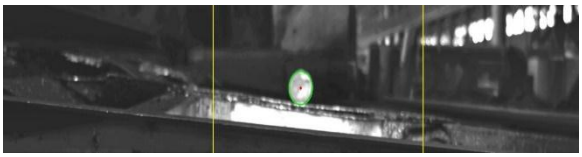
با بررسی سه روش یولو، تبدیل هاف دایره‌ای و تطبیق الگو، نتیجه‌گیری شد که تبدیل هاف دایره‌ای، بهترین توازن بین سرعت، دقت و سادگی تنظیم پارامترها را برای شرایط صنعتی موجود فراهم می‌کند. اگرچه یولو از نظر دقت بالقوه در تشخیص اشیاء پیچیده‌تر عملکرد بهتری دارد، اما نیاز به داده‌های آموزشی برچسب‌خورده فراوان و فرآیند آموزش زمان‌بر دارد که در محیط‌های صنعتی همیشه امکان‌پذیر نیست. از طرفی تطبیق الگو نیز به دلیل حساسیت بالا نسبت به تغییرات نوری و هندسی انتخاب مناسبی نبود. در مقابل، تبدیل هاف دایره‌ای بدون نیاز به آموزش مدل، امکان پیاده‌سازی سریع و تنظیم مستقیم پارامترهای آشکارسازی شعاع دایره را فراهم می‌کند.

علاوه بر این، الگوریتم هاف قابلیت اجرا روی سخت‌افزارهای موجود در خطوط تولید را بدون نیاز به واحد پردازش گرافیکی فراهم می‌سازد. در این پژوهش پس از انتخاب تبدیل هاف دایره‌ای، فرآیند استخراج مراکز سطح مقطع شمش‌ها با دقت و سرعت کافی انجام شد و داده‌های به‌دست آمده به مرحله تحلیل حرکت و استخراج ویژگی‌ها منتقل شدند.

٤-٣ یکنواخت‌سازی مسیر حرکتی

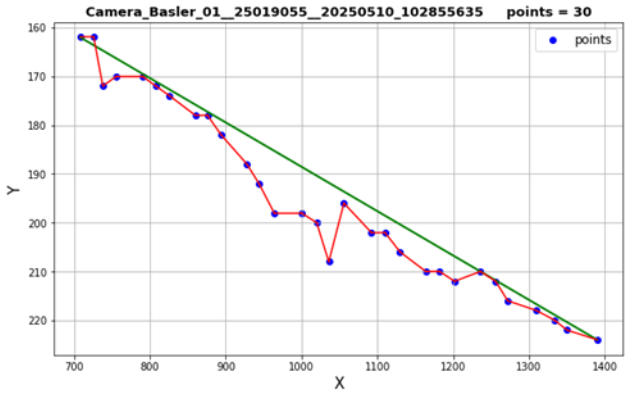
پس از بهبود کیفیت تصویر، خطوط مرزی (خط شروع و خط پایان) در فریم‌ها مشخص شدند تا تنها نقاطی که در بازه حرکتی مورد نظر قرار دارند، جمع‌آوری شوند. این کار باعث شد داده‌ها از نظر مکانی نیز استاندارد شوند و اثر ورود و خروج شمش به صحنه و نواحی غیرقابل اعتماد کاهش یابد. در هر فریم، مرکز سطح مقطع شمش تشخیص داده شد و به عنوان یک نقطه حرکتی در مسیر غلتیدن ثبت گردید.

در نهایت، برای یکسان‌سازی نمونه‌ها تعداد نقاط استخراج شده از هر ویدیو به مقدار مشخصی (مثلاً ۳۰ نقطه) تنظیم شد. اگر ویدیو تعداد بیشتری نقطه داشت، نقاط به‌طور یکنواخت از کل مسیر انتخاب شدند و اگر کمتر بود، از روش درون‌یابی استفاده شد که باعث شد ویژگی‌های حرکتی تنها منعکس‌کننده رفتار دینامیکی شمش باشند و نه صرفاً طول مسیر یا تعداد فریم‌ها. به بیان دیگر این کار مانع از آن شد که طول ویدیو یا تعداد فریم‌های ثبت شده بر نتایج تحلیلی تأثیر نامطلوب بگذارد. بدین ترتیب، مجموعه‌ای از نقاط حرکتی استاندارد و قابل مقایسه برای استخراج ویژگی‌ها به‌دست آمد. در شکل ۸ خطوط مرزی با رنگ زرد نشان داده شده است، در این فریم تشخیص سطح مقطع با روش تبدیل هاف دایره‌ای انجام شد و مرکز سطح مقطع شمش به عنوان یک نقطه حرکتی در مسیر غلتیدن ثبت گردید.

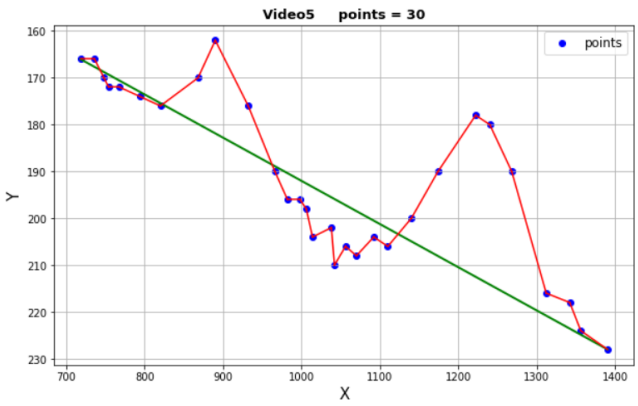


شکل (۸): تشخیص مرکز سطح مقطع شمش

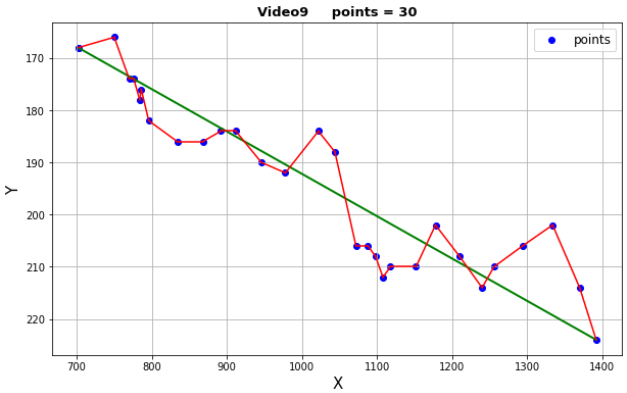
همچنین شکل ۹، تصویر مسیر حرکتی شمش سالم در ویدیو و شکل‌های ۱۰ تا ۱۲، به‌ترتیب تصویر مسیر حرکتی شمش‌های گرد با درجات متفاوت انحراف را نشان می‌دهند.



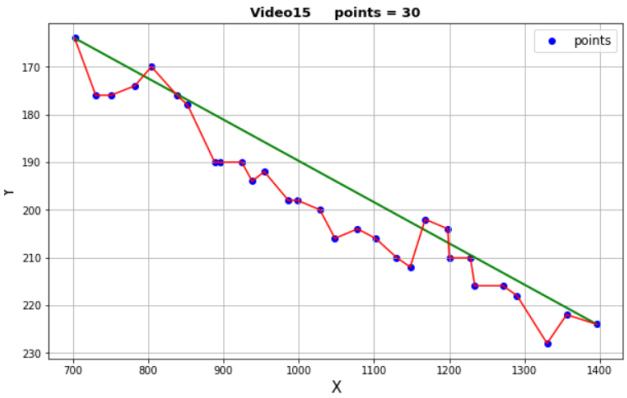
شکل (۹): مسیر حرکتی ترسیم شده از مرکز سطح مقطع‌های شمش سالم در طول خطوط مرزی حین غلتیدن روی میز خروجی



شکل (۱۰): مسیر حرکتی ترسیم شده از مرکز سطح مقطع‌های شمش خیلی ناصاف در طول خطوط مرزی حین غلتیدن روی میز خروجی



شکل (۱۱): مسیر حرکتی ترسیم شده از مرکز سطح مقطع‌های شمش ناصاف در طول خطوط مرزی حین غلتیدن روی میز خروجی



شکل (۱۲): مسیر حرکتی ترسیم شده از مرکز سطح مقطع‌های شمش کمی ناصاف در طول خطوط مرزی حین غلتیدن روی میز خروجی

٤-٤ استخراج ویژگی‌های حرکتی

در این پژوهش مجموعه‌ای از ویژگی‌های حرکتی استخراج شد که طراحی آن‌ها به گونه‌ای است که نه تنها برای تشخیص ناصافی شمش‌های فولادی کاربرد دارند، بلکه قابلیت استفاده در سایر سناریوهای صنعتی را نیز دارا هستند. نخستین ویژگی طول مسیر حرکتی[1] است که به‌عنوان مجموع فاصله‌های متوالی مراکز سطح مقطع در طول حرکت تعریف می‌شود. این شاخص میزان لغزش یا پیشروی مسیر را نشان می‌دهد و مستقل از جهت‌گیری دوربین قابل محاسبه است. از این رو، هر نوع حرکت دوبعدی جسم را می‌توان با این ویژگی ارزیابی کرد. این ویژگی از رابطه ۱ بدست می‌آید.

$$\mathrm{TL} = \sum_{i=1}^{N-1} \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2} \qquad (۱)$$

که در آن x_i و y_i مختصات مرکز سطح مقطع شمش در فریم iام و N تعداد فریم‌ها است.

ویژگی دوم مجموع مربعات انحراف[2] است. در این روش برای هر نقطه، فاصله عمودی آن از خط مستقیم بین نقطه شروع و پایان مسیر محاسبه شده و سپس مربع این فاصله‌ها با هم جمع می‌شود. این ویژگی معیاری برای سنجش میزان انحراف مسیر از حرکت ایده‌آل (خط راست) است و در مسیرهایی که نوسان یا تاب‌خوردگی بیشتری دارند، مقدار بالاتری خواهد داشت. نحوه محاسبه آن در رابطه ۲ نشان داده شده است.

$$\mathrm{SSD} = \sum_{i=1}^{N} d_i^2, \qquad d_i = \frac{|(x_N-x_1)(y_1-y_i)-(y_N-y_1)(x_1-x_i)|}{\sqrt{(x_N-x_1)^2+(y_N-y_1)^2}} \qquad (۲)$$

که در آن d_i فاصله عمودی هر نقطه از خط مرجع مسیر و (x_1, y_1) و (x_N, y_N) به‌ترتیب مختصات نقاط شروع و پایان مسیر هستند. نسخه ساده‌تر این ویژگی یعنی مجموع مربعات انحراف عمودی[3]، تنها در راستای عمودی مسیر محاسبه می‌شود و می‌تواند تغییرات در جهت اصلی ارتعاش را به‌طور خاص‌تر آشکار کند و مطابق رابطه ۳ محاسبه می‌شود.

$$\mathrm{SSVD} = \sum_{i=1}^{N} (y_i - {y^{ref}}_i)^2 \qquad (۳)$$

در این رابطه ${y^{ref}}_i$ مقدار متناظر روی خط بین نقاط ابتدایی و انتهایی مسیر برای مختصات x_i است.

برای بررسی نوسانات کوتاه‌مدت مسیر، ویژگی انحراف معیار ارتعاش[4] تعریف شد که بر اساس مشتق اول مختصات عمودی محاسبه می‌شود که در رابطه ٤ آمده است. این ویژگی توانایی بالایی در آشکارسازی لرزش‌های سریع و محلی دارد.

$$\mathrm{VSD} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(\Delta y_i - \overline{\Delta y})^2}, \qquad (٤)$$

[1] Trajectory Length
[2] Sum of Squared Deviations
[3] Sum of Squared Vertical Deviations
[4] Vibration of Standard Deviation

$$\Delta y_i = y_{i+1} - y_i, \qquad i = 1..N-1$$

در این رابطه $\overline{\Delta y}$ میانگین تغییرات عمودی در کل مسیر و Δy_i اختلاف موقعیت عمودی بین دو فریم متوالی است.

در ادامه، برای تحلیل دقیق‌تر رفتار ارتعاشی، از انرژی ارتعاش مبتنی بر تبدیل موجک[1] استفاده شد. در این روش انرژی جزئیات سطح اول سیگنال به کمک تبدیل موجک محاسبه می‌شود که امکان تمرکز بر اجزای پر فرکانس ارتعاش را فراهم می‌سازد. نحوه محاسبه آن در رابطه ۵ آمده است.

$$\mathrm{WVE} = \frac{1}{K}\sum_{k=1}^{K}(d_k^{(1)})^2 \qquad (۵)$$

که در آن $d_k^{(1)}$ ضرایب جزئیات سطح اول تبدیل موجک و K تعداد ضرایب جزئیات است.

ویژگی مهم دیگر انحراف معیار شتاب[2] است که بر اساس تغییرات سرعت محاسبه می‌شود. مقدار بزرگ‌تر این ویژگی بیانگر نوسانات شدیدتر در دینامیک حرکت است و می‌تواند برای تمایز نمونه‌های سالم و ناصاف کاربرد بالایی داشته باشد. نحوه محاسبه این ویژگی در رابطه ٦ نشان داده شده است.

$$\mathrm{ASD} = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N-2}(a_i - \bar{a})^2}, \qquad (٦)$$

$$a_i = v_{i+1} - v_i,$$

$$v_i = \sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2}$$

که در آن $\bar{a}$ میانگین شتاب در کل مسیر، a_i تغییرات سرعت (یعنی شتاب نسبی) و v_i سرعت لحظه‌ای بین دو فریم است.

لازم به ذکر است که همه‌ی محاسبات بر روی مسیر یکنواخت‌شده با $N = M$ نمونه انجام می‌شود تا تأثیر طول مسیر و تعداد فریم حذف گردد. در این پژوهش مقدار M برابر ۳۰ تعیین شد. در نهایت، برای جلوگیری از تأثیر مقیاس‌های متفاوت این ویژگی‌ها بر روند تحلیل، تمامی داده‌ها با استفاده از روش Z-score [۱۳] نرمال‌سازی شدند. این کار باعث می‌شود مقایسه ویژگی‌ها در یک فضای مشترک انجام شود و انتخاب آستانه‌ها یا معیارهای طبقه‌بندی با کمترین سوگیری ناشی از مقیاس صورت گیرد.

٤-٥ روش‌های تصمیم‌گیری

پس از استخراج ویژگی‌های حرکتی که بیانگر رفتار شمش در هنگام غلتیدن هستند، گام بعدی انتخاب روشی مناسب برای تبدیل این ویژگی‌ها به یک طبقه‌بندی دودویی (سالم و ناصاف) است. برای این منظور، پس از بررسی روش‌های مختلف سه روش در میان آن‌ها انتخاب شد و در آن‌ها شمش‌های ناصاف به‌عنوان کلاس مثبت و شمش‌های سالم به‌عنوان کلاس منفی در نظر گرفته شدند.

روش اول، آستانه‌گذاری تک‌مرحله‌ای است. در این روش، برای هر ویژگی یک آستانه بهینه تعیین می‌شود تا داده‌ها به دو دسته جداگانه تقسیم شوند. انتخاب این آستانه با استفاده از نمودار مشخصه عملکرد سیستم انجام می‌شود. این منحنی با ترسیم نرخ مثبت درست[3] در برابر نرخ مثبت نادرست[4] ساخته می‌شود. نرخ مثبت درست که با عنوان حساسیت[5] نیز شناخته می‌شود، نسبت نمونه‌های مثبت به درستی شناسایی‌شده را نشان می‌دهد. در مقابل، نرخ مثبت نادرست احتمال شناسایی نادرست نمونه‌های منفی به‌عنوان مثبت را نشان می‌دهد. از آنجا که نمودار مشخصه عملکرد سیستم تمامی آستانه‌های ممکن را ارزیابی می‌کند، انتخاب آستانه به‌گونه‌ای صورت می‌گیرد که اختلاف بین نرخ مثبت درست و نرخ مثبت نادرست بیشینه شود و بدین ترتیب بهترین نقطه برای جداسازی دو کلاس مشخص گردد.

روش دوم، آستانه‌گذاری چندمرحله‌ای است که با توجه به حساسیت بیشتر نسبت به هزینه مسئله طراحی شده است. در مسئله حاضر، کلاس شمش ناصاف اهمیت ویژه‌ای دارد، زیرا خطای تشخیص شمش ناصاف به‌عنوان سالم بسیار پرهزینه‌تر از خطای معکوس آن خواهد بود. بنابراین، در طراحی این رویکرد توجه ویژه‌ای به کاهش موارد منفی نادرست شده است، حتی اگر این امر باعث افزایش اندک مثبت‌های نادرست گردد. معیارهای اصلی در این روش شامل سه اصل است:

۱- حداکثرسازی مثبت‌های درست و منفی‌های درست

[1] Wavelet Vibration Energy
[2] Acceleration's Standard Deviation
[3] True Positive Rate (TPR)
[4] False Positive Rate (FPR)
[5] Recall

۲- به حداقل رساندن منفی‌های نادرست به‌عنوان مهم‌ترین اولویت
۳- کاهش مثبت‌های نادرست با اهمیت ثانویه.

این ساختار چندمرحله‌ای عملاً یک غربال‌گری تدریجی ایجاد می‌کند تا موارد مرزی با دقت بیشتری بررسی شوند و خطر از دست رفتن یک نمونه ناصاف به حداقل برسد.

در پیاده‌سازی عملی این روش، برای هر ویژگی، نمودارهای دقت پیش‌بینی- حساسیت[1] و مشخصه عملکرد سیستم محاسبه گردید تا عملکرد الگوریتم در سطوح مختلف آستانه بررسی شود. در گام نخست، آستانه‌هایی انتخاب شدند که سه شرط زیر را هم‌زمان برآورده کنند:

۱. مقدار حساسیت حداقل برابر با ۰/۹ باشد تا اطمینان حاصل شود که حداقل ۹۰ درصد از شمش‌های ناصاف به‌درستی شناسایی می‌شوند.
۲. مقدار دقت پیش‌بینی[2]، کمتر از ۰/٦ نباشد تا میزان هشدارهای اشتباه در حد قابل‌قبول باقی بماند.
۳. تعداد منفی‌های درست بیش از نیمی از کل شمش‌های سالم باشد تا صحت شناسایی محصولات سالم نیز حفظ شود.

در میان آستانه‌هایی که این سه شرط را برآورده می‌کردند، مقداری به‌عنوان آستانه بهینه انتخاب شد که بیشترین مقدار شاخص F_2-Score را داشت. در صورتی که هیچ آستانه‌ای تمامی شرایط را برآورده نمی‌کرد، با حذف شرط دوم، آستانه‌ای با بیشترین دقت[3] انتخاب می‌گردید. اگر در این مرحله نیز آستانه‌ای یافت نمی‌شد، آستانه‌ای با بیشترین مقدار حساسیت در عین حفظ تعداد منفی‌های درست بالا برگزیده می‌شد. در نهایت، در حالتی که هیچ‌کدام از شرایط قبلی برآورده نمی‌شدند، آستانه‌ای انتخاب می‌گردید که بیشترین توان تفکیک بین دو کلاس را بر اساس اختلاف میان نرخ مثبت درست و مثبت نادرست ایجاد کند. این منطق چندمرحله‌ای موجب شد تا فرآیند تصمیم‌گیری با تمرکز بر کاهش موارد منفی نادرست انجام گیرد، حتی اگر این امر موجب افزایش اندکی در خطاهای مثبت نادرست شود.

روش سوم، خوشه‌بندی مبتنی بر الگوریتم K-Means است که براساس مقادیر ویژگی‌های غالب، یکی از خوشه‌ها به‌عنوان نماینده شمش‌های سالم و دیگری به‌عنوان نماینده شمش‌های ناصاف برچسب‌گذاری شد. در این روش، تعیین کلاس برای داده‌های جدید نیز در واقع براساس مقایسه فاصله اقلیدسی نمونه با مراکز خوشه‌ها انجام می‌شود.

در نهایت، هر یک از این سه روش به‌صورت جداگانه بر روی داده‌های استخراج‌شده اعمال شدند تا کارایی و قابلیت تفکیک آن‌ها در شرایط صنعتی بررسی شود. جزئیات نتایج و تحلیل مقایسه‌ای عملکرد این رویکردها در بخش پنجم ارائه شده است. همچنین، نمای کلی از شبه‌کد روش پیشنهادی در شکل ۱۳ نشان داده شده است که به‌صورت مرحله‌به‌مرحله تمام مراحل پردازش، استخراج ویژگی و تصمیم‌گیری را نمایش می‌دهد.

```
Input: video, lines A/B, target_points = M
Output: label ∈ {Healthy, Defective}, features

1. video ← enhance(video; contrast, CLAHE); assert
fps≈100
2. for each frame:
     detect circle center c = (cx, cy) using Hough
     if A ≤ cx ≤ B: centers.append(c)
3. centers ← uniform_resample(centers, M)  #
interpolation if needed
4. x ← centers[:,0]; y ← centers[:,1]

5. Extract Features:
  trajectory_length     = Σ ||c[i+1]-c[i]||
  sum_sq_dev      = Σ d(point, line(start,end))^2
  sum_vertical_sq_dev       = Σ (y[i] - line_y[i])^2
  vibration_std  = std(diff(y))
  ay                = diff(diff(y))
  wavelet_vib_energy     = Σ (DWT_details(ay))^2 / |ay|
  acceleration_std.   = std(diff(√(diff(x)^2 + diff(y)^2)))

6. features ← Zscore(features)

7. Decision Methods:
  (a) Single-threshold:
     τ ← argmax_τ { TPR(τ) - FPR(τ) } (Youden's J via
ROC)
     label = Bent if feature* > τ else Straight

  (b) Multi-stage Thresholding:
     For each feature fi:
       Evaluate candidate thresholds by Recall, Precision,
TN, Fβ (β=2)
       Priority:
         Step 1 → max(Fβ) if Recall≥0.90 ∧
Precision≥0.60 ∧ TN≥50%
         Step 2 → max(TP+TN) if Recall≥0.90 ∧ TN≥50%
         Step 3 → max(Recall) under TN≥50%
         Step 4 → fallback to Youden's J
       label = Bent if fi > τi else Straight

  (c) K-Means clustering:
     X ← all_features_normalized
     clusters ← kmeans(X, k=2)
     assign cluster with higher Bent ratio → label=Bent
     label = cluster(video_features)

8. Report:
  Confusion Matrix, Precision, Recall, F1-score,
ROC/AUC, PCA-2D clusters
```

شکل (۱۳): شبه‌کد روش پیشنهادی

[1] Precision–Recall

[2] Precision
[3] Accuracy

۵- ارزیابی نتایج

در مراحل اولیه، ویدیوهای متعددی از فرآیند غلتیدن شمش‌های خروجی دستگاه صافکاری در کارخانه فولاد آلیاژی ایران ثبت شد، اما همه‌ی آن‌ها از کیفیت لازم برای تحلیل برخوردار نبودند. رفتار شمش‌ها روی میز خروجی همیشه قابل پیش‌بینی نیست؛ برخی شمش‌ها به‌دلیل ناصافی یا وضعیت سطح میز، در مسیر حرکت کج شده یا حتی متوقف می‌شوند و این موارد باعث حذف بخشی از داده‌ها شد. از سوی دیگر، فرآیند ضبط هر ویدیو نیازمند هماهنگی میان چندین واحد کارخانه، توقف موقت خط تولید و تنظیم دقیق پارامترهای نور، زاویه دید و شاتر دوربین باسلر بود. در هر نوبت آزمایش، تنظیمات تصویربرداری باید مجدداً انجام می‌شد تا وضوح و ثبات تصویر تضمین شود. در نهایت، پس از تلاش و انجام چندین مرحله آزمایشی، مجموعه‌ای شامل ۲۳ ویدیو با کیفیت بالا و شرایط کنترل‌شده تهیه گردید که هرکدام نمایانگر حرکت یک شمش خروجی بر روی میز شیب‌دار دستگاه صافکاری هستند. از میان این نمونه‌ها، ٦ شمش توسط کارشناسان به‌عنوان سالم[1] و ۱۷ شمش به‌عنوان ناصاف[2] برچسب‌گذاری شدند. اگرچه تعداد نمونه‌ها محدود است، اما تمامی آن‌ها در شرایط واقعی صنعتی و با کیفیت یکنواخت ثبت شده‌اند تا ارزیابی روش پیشنهادی از اعتبار فنی کافی برخوردار باشد.

۵-۱ معیارهای ارزیابی

برای ارزیابی عملکرد روش پیشنهادی نیاز به معیارهایی استاندارد و قابل اعتماد است. در این راستا، از معیارهایی نظیر دقت، حساسیت، دقت پیش‌بینی و امتیاز F1 استفاده شد که نحوه محاسبه آن‌ها در رابطه‌های ۷ تا ۱۰ بیان شده است.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۷)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (۸)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۹)$$

$$F1 = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 \quad (۱۰)$$

در رابطه‌های ۷ تا ۱۰، TP مثبت درست است، یعنی تعداد برچسب های مثبتی که به درستی توسط طبقه‌بندی‌کننده مثبت پیش‌بینی شده است. TN منفی درست است، تعداد برچسب‌های منفی که توسط طبقه‌بندی‌کننده به درستی منفی پیش‌بینی شده است. FP مثبت نادرست، تعداد برچسب‌های منفی که توسط طبقه‌بندی‌کننده به اشتباه مثبت پیش‌بینی شده است و FN نشان دهنده منفی نادرست است و به تعداد برچسب‌های مثبت که توسط طبقه‌بندی‌کننده به اشتباه منفی پیش‌بینی شده است اشاره می‌کند.

۵-۲ تحلیل ویژگی‌ها با نمودار جعبه‌ای

در این مطالعه، نمودار جعبه‌ای به‌عنوان ابزاری کلیدی برای مقایسه توزیع مقادیر ویژگی‌ها در میان دو گروه سالم و ناصاف به‌کار گرفته شد. این نمودار امکان نمایش شاخص‌های مرکزی مانند میانه، چارک‌ها و همچنین شناسایی مقادیر پرت را فراهم می‌آورد و نشان می‌دهد کدام ویژگی‌ها قدرت تمایز بیشتری دارند. این نمودار برای هر ویژگی ترسیم شد و نتایج نشان داد که شمش‌های صاف و ناصاف همپوشانی دارند ولی بهترین تمایزها مربوط به ویژگی‌های مجموع مربعات انحراف، مجموع مربعات انحراف عمودی و انحراف معیار ارتعاش هستند.

۵-۳ نتایج روش‌های تصمیم‌گیری

پس از تحلیل اولیه ویژگی‌ها، سه روش تصمیم‌گیری معرفی‌شده در بخش ٤-۵ بر روی ویژگی‌ها آزمایش شدند و دقت و پایداری هر یک از روش‌ها در شرایط صنعتی بررسی شد.

برای ارزیابی توان پردازشی، زمان اجرای بخش‌های کلیدی سیستم روی یک رایانه با مشخصات سخت‌افزاری (CPU: Intel i7 / RAM: 16GB) اندازه‌گیری شد. میانگین زمان پردازش هر فریم برای پیش‌پردازش حدود ۵ ms، برای تبدیل هاف دایره‌ای حدود ۱۱ ms و برای استخراج ویژگی‌ها و تصمیم‌گیری نهایی کمتر از ۱ ms بوده است. در مجموع، میانگین زمان پردازش هر شمش زیر یک ثانیه است. این سرعت نشان می‌دهد که الگوریتم

[1] Healthy

[2] Defective

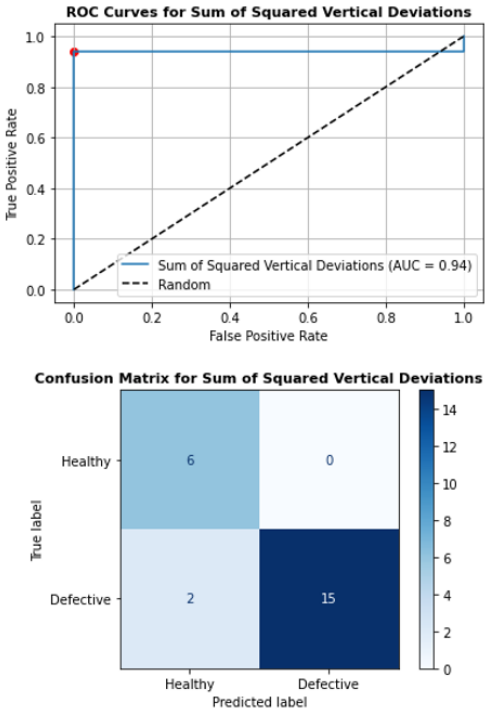
پیشنهادی بدون نیاز به سخت‌افزار خاص، توانایی اجرا در زمان واقعی در محیط صنعتی را دارد.

در فرآیند عملی، فاصله زمانی مشخص حداقل ۴۰ ثانیه میان خروج هر شمش از دستگاه صافکاری و ورود شمش بعدی به روی میز وجود دارد. با توجه به این بازه، سیستم فرصت کافی دارد تا ویدیو را پردازش کرده، ویژگی‌ها را استخراج نموده و تصمیم نهایی درباره سالم یا ناصاف بودن شمش را اعلام کند. در چنین حالتی، به محض تشخیص ناصافی، سیستم می‌تواند به اپراتور هشدار دهد یا به‌صورت خودکار مسیر خروج شمش را تغییر دهد تا از ورود محصول معیوب به خطوط بعدی جلوگیری شود.

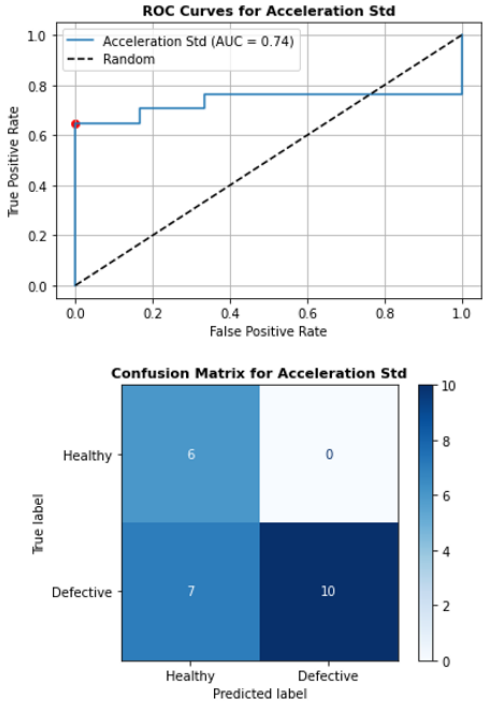
در ادامه، در این بخش ابتدا نتایج روش آستانه‌گذاری تک‌مرحله‌ای، سپس آستانه‌گذاری چندمرحله‌ای و در نهایت روش خوشه‌بندی مبتنی بر K-Means ارائه و مقایسه شدند.

نتایج روش آستانه‌گذاری تک‌مرحله‌ای نشان داد که بهترین عملکرد به‌ترتیب مربوط به ویژگی‌های مجموع مربعات انحراف عمودی، انحراف معیار ارتعاش و مجموع مربعات انحراف هستند. نمونه‌ای از این نتایج در شکل‌های ۱۴ و ۱۵ آمده است. به‌طور خاص، ویژگی مجموع مربعات انحراف عمودی با آستانه منفی ۰/۵۷ در شکل ۱۵ با سطح زیر منحنی ۰/۹۴ توانسته است عملکرد بسیار خوبی در جداسازی شمش‌های سالم و ناصاف ارائه دهد. همچنین، ماتریس درهم‌ریختگی این ویژگی نشان می‌دهد که از ۱۷ شمش ناصاف، ۱۵ مورد به‌درستی شناسایی شده و تنها ۲ مورد خطا داشته است، در حالی که تمامی ۶ شمش سالم بدون خطا تشخیص داده شده‌اند. این نتایج بیانگر قدرت بالای این ویژگی برای کاربردهای بازرسی صنعتی است.

روش آستانه‌گذاری چندمرحله‌ای با توجه به هزینه‌های متفاوت خطا طراحی شد و نتایج نشان داد که سه ویژگی انحراف معیار ارتعاش، مجموع مربعات انحراف عمودی و مجموع مربعات انحراف بیشترین قدرت تفکیک را داشتند. ماتریس‌های درهم‌ریختگی برای هر ویژگی محاسبه شد و نمونه‌هایی از این نتایج در شکل‌های ۱۶ و ۱۷ نمایش داده شده است که بیانگر بهبود در شناسایی شمش‌های ناصاف هستند.

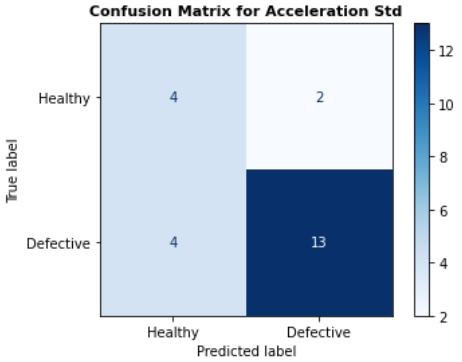


شکل (۱۴): نمودار مشخصه عملکرد سیستم و ماتریس درهم‌ریختگی برای ویژگی مجموع مربعات انحراف عمودی

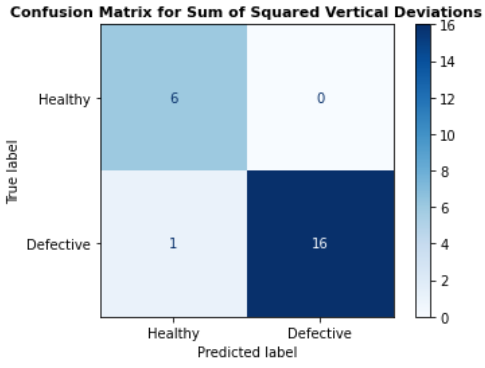


شکل (۱۵): نمودار مشخصه عملکرد سیستم و ماتریس درهم‌ریختگی برای ویژگی انحراف معیار شتاب

تحلیل دقیق‌تر نتایج نشان می‌دهد که این استراتژی توانست خطای منفی نادرست را کاهش دهد. برای مثال، در ویژگی انحراف معیار شتاب با آستانه منفی ۰/۷۲ تعداد خطاهای منفی نادرست از ۷ در روش تک‌مرحله‌ای به ٤ کاهش یافت، به این معنا که تعداد شمش‌های ناصاف شناسایی‌شده از ۱۰ به ۱۳ افزایش پیدا کرد. هرچند که در این روش تعداد خطاهای مثبت نادرست از صفر به ۲ مورد افزایش یافت، اما با توجه به اهمیت بالاتر جلوگیری از رد شدن شمش‌های ناصاف، این تغییر یک بهبود ارزشمند محسوب می‌شود. با این حال کاهش خطاهای منفی نادرست با هزینه افزایش خطاهای مثبت نادرست در تمام ویژگی‌ها رخ نداد و ما در ویژگی مجموع مربعات انحراف عمودی با آستانه منفی ۰/۸۵، شاهد بهبود چشم‌گیری بوده‌ایم، به‌طوری‌که خطای منفی نادرست از ۲ به ۱ کاهش پیدا کرد و در عین حال هیچ افزایشی در خطاهای مثبت نادرست مشاهده نشد. این نتیجه نشان می‌دهد که این ویژگی نه تنها در شناسایی شمش‌های ناصاف موفق عمل کرده، بلکه توانسته تمامی نمونه‌های سالم را نیز به درستی تشخیص دهد.



شکل (۱٦): ماتریس درهم‌ریختگی ویژگی انحراف معیار شتاب

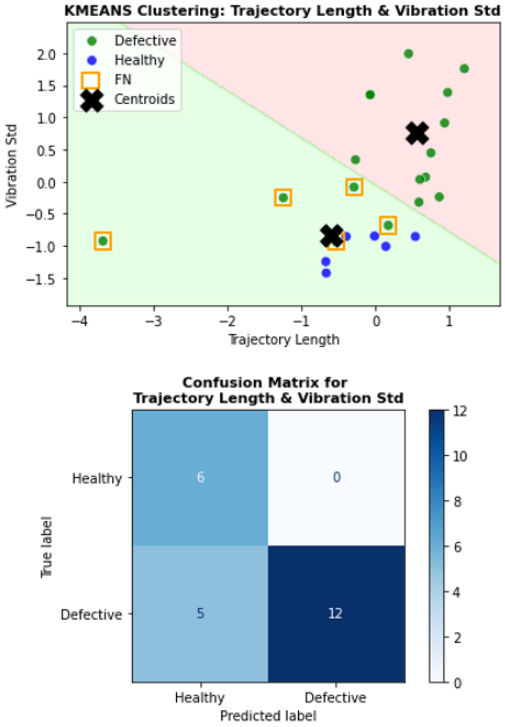


شکل (۱۷): ماتریس درهم‌ریختگی ویژگی مجموع مربعات انحراف عمودی

در بررسی تأثیر مقدار آستانه، مشاهده شد که جهت تغییرات عملکرد به نوع ویژگی بستگی دارد. در ویژگی‌هایی مانند مجموع مربعات انحراف که مقدار آن با ناصافی شمش افزایش می‌یابد، با افزایش آستانه نمونه‌های ناصاف کمتری شناسایی شده و احتمال از دست رفتن آن‌ها افزایش می‌یابد، در نتیجه کاهش حساسیت کاهش و دقت پیش‌بینی افزایش می‌یابد. در مقابل، در ویژگی‌هایی نظیر انحراف معیار شتاب که رفتار معکوس دارند، کاهش آستانه همان اثر را ایجاد می‌کند.

در نهایت، نتایج روش خوشه‌بندی نشان داد که این روش توانسته است تا حد زیادی داده‌ها را بر اساس ساختار فاصله‌ای خوشه‌بندی کند، اما به دلیل محدودیت ذاتی این روش در استفاده از مرزهای خطی، توانایی کافی برای تشخیص داده‌های پیچیده یا هم‌پوشان را ندارد. در این خروجی، داده‌های نزدیک به مرز بیشترین میزان خطا را داشته‌اند، که امری طبیعی در روش‌های مبتنی بر فاصله است.

در ادامه، نتیجه روش خوشه‌بندی روی دو ویژگی طول مسیر حرکتی و انحراف معیار ارتعاش در شکل ۱۸ نمایش داده شده است.



شکل (۱۸): نتیجه روش خوشه‌بندی روی دو ویژگی طول مسیر حرکتی و انحراف معیار ارتعاش

همانطور که مشاهده می‌شود، مراکز خوشه‌ها به‌وسیله الگوریتم مشخص شده‌اند و نواحی تصمیم‌گیری بر اساس فاصله اقلیدسی شکل گرفته‌است. نقاط سبز نشان‌دهنده شمش‌های ناصاف و نقاط آبی نمایانگر شمش‌های سالم هستند که این امکان را فراهم می‌کند تا عملکرد الگوریتم در مقایسه با واقعیت ارزیابی شود. با این حال، داده‌های نزدیک به مرزها بیشترین میزان خطا را داشتند. همچنین ماتریس درهم‌ریختگی نشان می‌دهد که تمامی نمونه‌های سالم به‌درستی شناسایی شده‌اند، اما پنج نمونه از شمش‌های ناصاف به اشتباه در گروه سالم قرار گرفتند. این موضوع باعث شد دقت کلی مدل حدود ۷۸ درصد باشد.

٦- نتیجه گیری و کارهای آینده

در این پژوهش، روشی مبتنی بر تحلیل مسیر حرکت سطح مقطع شمش‌های فولادی جهت تشخیص خودکار ناصافی ارائه شد. نتایج آزمایش‌ها نشان داد که ویژگی‌هایی مانند مجموع مربعات انحراف عمودی و انحراف معیار ارتعاش بیشترین قدرت تمایز را میان شمش‌های سالم و ناصاف دارند. مقایسه‌ی روش‌های تصمیم‌گیری نشان داد که رویکرد آستانه‌گذاری چندمرحله‌ای با درنظر گرفتن هزینه‌های نامتقارن خطا، عملکرد دقیق‌تر و باثبات‌تری نسبت به سایر روش‌ها ارائه می‌دهد. این منطق چندمرحله‌ای باعث شد سیستم ضمن افزایش حساسیت نسبت به شناسایی شمش‌های ناصاف، ثبات خود را در حفظ دقت تشخیص نمونه‌های سالم نیز حفظ کند و در نتیجه توازنی میان ایمنی تشخیص (حساسیت بالا) و قابلیت اطمینان صنعتی (دقت پیش‌بینی قابل قبول) برقرار شود. ارزیابی زمان اجرای الگوریتم نشان داد که میانگین زمان پردازش هر شمش زیر یک ثانیه است؛ بنابراین روش پیشنهادی بدون نیاز به سخت‌افزار خاص، قابلیت اجرا در زمان واقعی را دارد و می‌تواند در بازه زمانی بین عبور دو شمش روی میز خروجی تصمیم نهایی را اعلام کند. هرچند تعداد نمونه‌های مورد استفاده به دلیل محدودیت‌های عملیاتی و الزامات ایمنی کارخانه محدود بوده است، اما داده‌ها از تنوع حرکتی کافی برخوردار بودند و برای ارزیابی هدف پژوهش کفایت دارند. در ادامه‌ی این پژوهش، برنامه جمع‌آوری داده‌های گسترده‌تر، بهبود شرایط نوری و استفاده از چند موقعیت دوربین در نظر گرفته شده است تا دقت و تعمیم‌پذیری مدل افزایش یابد.

References

[1] Y. J. Chen, Y. H. Yeh, and J. F. Yang, “A system for the real-time detection of the U-shaped steel bar straightness on a production line,” Sensors, vol. 25, no. 13, p. 3972, 2025

[2] Y. Saif, A. Z. M. Rus, Y. Yusof, M. L. Ahmed, S. Al-Alimi, D. H. Didane, and H. Q. A. Abdulrab, “Advancements in roundness measurement parts for industrial automation using Internet of Things architecture-based computer vision and image processing techniques,” Applied Sciences, vol. 13, no. 20, p. 11419, 2023.

[3] O. Sidla, E. Wilding, A. Niel, and H. Barg, “Vision system for gauging and automatic straightening of steel bars,” in Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology, vol. 4189, pp. 248–257, Feb. 2001.

[4] X. Chen, M. Bi, and C. Xue, “Study on straightness measurement technology of slender rod based on time-domain three-point method,” in J. Phys.: Conf. Ser., vol. 1303, no. 1, p. 012140, Aug. 2019, IOP Publishing.

[5] T. Schmid-Schirling, L. Kraft, and D. Carl, “Laser scanning–based straightness measurement of precision bright steel rods at one point,” Int. J. Adv. Manuf. Technol., vol. 116, no. 7, pp. 2511–2519, 2021.

[6] Y. Jiang, “Surface defect detection of steel based on improved YOLOv5 algorithm,” Electronic Research Archive, vol. 31, no. 11, 2023.

[7] R. Usamentiaga and D. F. Garcia, “Rail flatness measurement based on dual laser triangulation,” IEEE Trans. Ind. Appl., vol. 60, no. 4, pp. 5849–5860, 2024.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 779–788.

[9] C. H. Kang and S. Y. Kim, “Real-time object detection and segmentation technology: An analysis of the YOLO algorithm,” JMST Adv., vol. 5, no. 2, pp. 69–76, 2023.

[10] N. M. Alahdal, F. Abukhodair, L. H. Meftah, and A. Cherif, “Real-time object detection in autonomous vehicles with YOLO,” Procedia Comput. Sci., vol. 246, pp. 2792–2801, 2024.

[11] Y. Xie and Q. Ji, “A new efficient ellipse detection method,” in Proc. 16th Int. Conf. Pattern Recognit., vol. 2, 2002, IEEE.

[12] R. Brunelli and T. Poggiot, “Template matching: Matched spatial filters and beyond,” Pattern Recognit., vol. 30, no. 5, pp. 751–768, 1997.

[13] S. W. Huck, T. L. Cross, and S. B. Clark, “Overcoming misconceptions about Z-scores,” Teaching Statistics, vol. 8, pp. 38–40, 1986.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 107-123

Automatic Detection of Defective Round Steel Billets from Straightening Machine Output Using Machine Vision Techniques

Rastin Namiranian[1], Sahar Soltani[2], Vali Derhami[3*]

[1] Information Technology Department, Iran Alloy Steel Company, Yazd, Iran

[2]M.Sc. in Artificial Intelligence and Robotics, Iranian Alloy Steel Development Company, Yazd, Iran

[3] Professor, Department of Computer Engineering, Yazd University, Yazd, Iran

Article Information

Original Research Paper

Received:
2025 September 25

Accepted:
2025 November 14

Keywords:
Circular Hough Transform, machine Vison, Steel Billet, Straightening Machine

Corresponding Author*:
vderhami@yazd.ac.ir

Abstract

This paper addresses the problem of detecting defective round steel billets during their rolling motion. In the finishing hall of the Iran Alloy Steel Factory, one of the essential operations applied to round billets is straightening. However, no system currently exists to automatically assess the quality of the straightening machine's output. To address this challenge, videos of billets rolling on the discharge table after leaving the straightening machine are recorded. After preprocessing and illumination balancing, the cross-section of each billet in every frame is detected using the Circular Hough Transform method. The motion trajectory of the billet in each video is then determined, and six motion-related features are extracted for analysis. These features include Trajectory Length, Sum of Squared Deviations, Sum of Squared Vertical Deviations, Vibration of Standard Deviation, Wavelet Vibration Energy, and Standard Deviation of Acceleration. The values of these features are computed with normalization based on sampling per unit length. Experimental results demonstrate that features such as Vibration of Standard Deviation, Sum of Squared Vertical Deviations, and Sum of Squared Deviations show strong discriminative power in separating healthy billets from defective ones. Although the false negative rate is somewhat high, the proposed method achieves an overall accuracy of about 85%, which is considered acceptable for practical industrial application.

: 10.22034/ABMIR.2025.23720.1171

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۳۹-۱۲۵

SSLA-Net: معماری یادگیری معنایی ـ مکانی برای شناسایی مقاوم نقص‌های سطحی فولاد

معصومه رضائی[1*]، منصوره رضائی[2]، مهری رجائی[3]

[1] استادیار گروه مهندسی کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران

[2] دکتری هوش مصنوعی، دانشکده مهندسی کامپیوتر، پردیس فنی و مهندسی، دانشگاه یزد، یزد، ایران

[3] استادیارگروه مهندسی فناوری اطلاعات، دانشگاه سیستان و بلوچستان، زاهدان، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۶/۲۲

تاریخ پذیرش:
۱۴۰۴/۰۸/۱۲

کلیدواژه‌ها:
یادگیری چندوجهی، شناسایی نقص‌های سطحی، توجه چندسری، نمایش معنایی–مکانی، بینایی ماشین صنعتی

نویسنده‌مسئول:
mrezaei@ece.usb.ac.ir

چکیده

در صنایع فولاد، شناسایی دقیق و سریع نقص‌های سطحی نقشی حیاتی در تضمین کیفیت و کاهش ضایعات ایفا می‌کند. بااین‌حال، شباهت ظاهری میان الگوهای بافتی و پیچیدگی نواحی معیوب، کارایی روش‌های متداول بینایی ماشین را با محدودیت مواجه کرده است. در این پژوهش، یک چارچوب ترکیبی چندوجهی برای تشخیص و طبقه‌بندی نقص‌های سطحی فولاد ارائه شده است که با تلفیق مدل تشخیص ناحیه‌ای **Faster R-CNN** و نمایش‌های معنایی مدل **CLIP** از طریق ماژول توجه چندسری، ارتباط میان ویژگی‌های بصری و مفهومی را به‌صورت هم‌زمان مدل‌سازی می‌کند. این هم‌افزایی موجب افزایش دقت در شناسایی نواحی کوچک و پراکنده و همچنین بهبود تفکیک‌پذیری کلاس‌ها در فضای ویژگی شده است. در طراحی مدل، سه تابع ضرر مکمل شامل ضرر تشخیص ناحیه‌ای، ضرر هم‌ترازی ناحیه–متن و ضرر سطح تصویر به‌کار رفته‌اند تا یادگیری هم‌زمان سطوح مکانی و معنایی امکان‌پذیر گردد. نتایج تجربی روی مجموعه‌داده‌ی استاندارد **NEU-DET** نشان می‌دهند که مدل پیشنهادی به دقت ۱۰۰٪ در طبقه‌بندی سطح تصویر و **mAP@50**=۷۸/۰٤٪ در سطح ناحیه‌ای دست یافته است. این عملکرد برجسته نشان می‌دهد که یادگیری چندوجهی مبتنی بر **CLIP** می‌تواند دقت و پایداری تشخیص نقص را در شرایط واقعی صنعتی به‌طور چشمگیری ارتقا دهد.

: 10.22034/ABMIR.2025.23624.1166

E-ISSN: 2821-2037

۱- مقدمه

فولاد به‌عنوان یک ماده حیاتی در صنایع استراتژیک از جمله خودروسازی، هوافضا، ساخت‌وساز و تولید تجهیزات صنعتی شناخته می‌شود و نقشی کلیدی در کیفیت، ایمنی و دوام محصولات نهایی دارد. وجود هرگونه نقص سطحی نظیر ترک، خراش، پوسته شدگی یا فرورفتگی، منجر به کاهش استحکام مکانیکی، کاهش عمر مفید و بروز خرابی‌های جدی در خطوط تولید می‌گردد. ازاین‌رو، تشخیص سریع و دقیق نقص‌های سطح فولاد از الزامات فنی، اقتصادی و ایمنی محسوب می‌شود [۱]. علی‌رغم پیشرفت‌های اخیر در حوزه تشخیص خودکار عیوب سطحی، شناسایی نقص‌های کوچک، پراکنده و دارای بافت پیچیده همچنان چالشی اساسی محسوب می‌شود. این موضوع به‌ویژه در خطوط تولید صنعتی که نیازمند عملکرد پایدار و مقاوم هستند اهمیت بیشتری می‌یابد. به‌منظور پاسخ‌گویی به این نیازها، در این پژوهش یک چارچوب یادگیری چندوجهی به نام SSLA-Net پیشنهاد شده است که با ادغام اطلاعات مکانی و معنایی، توانایی مدل در تفکیک دقیق کلاس‌ها و شناسایی نقص‌های دشوار را افزایش می‌دهد.

روش‌های سنتی تشخیص نقص، شامل بازرسی چشمی و روش‌های غیرمخرب فیزیکی و شیمیایی (مانند التراسونیک و ذرات مغناطیسی)، در محیط‌های صنعتی مدرن با محدودیت‌هایی مواجه هستند. این روش‌ها عمدتاً زمان‌بر، پرهزینه و دارای توانایی محدود در شناسایی نقص‌های کوچک و مبهم هستند [۲]. در نتیجه، تحقیقات به سمت کاربرد یادگیری عمیق و بینایی ماشین معطوف شده است. روش‌های مبتنی بر شبکه‌های کانولوشنی (CNN[1])، ترنسفورمرها و معماری‌های ترکیبی، پیشرفت‌های قابل‌توجهی در تشخیص و طبقه‌بندی نقص‌ها در مجموعه‌داده‌های استاندارد مانند **NEU-DET** [۳] و GC10-DET [٤] نشان داده‌اند [۸-۵]. بااین‌حال، چالش‌هایی همچنان در زمینه تشخیص نقص‌های کوچک و پیچیده، مرزبندی دقیق کلاس‌های مشابه، و ایجاد نمایش‌های معنایی قوی برای نواحی تصویر وجود دارد.

به‌منظور رفع محدودیت‌های مذکور، این مطالعه یک چارچوب تلفیقی نوآورانه را ارائه می‌دهد. این چارچوب بر اساس ترکیب Faster R-CNN برای قابلیت‌های ناحیه محور، CLIP برای بازنمایی معنایی، و مکانیزم توجه چندسر[2]طراحی شده است. در این روش، علاوه بر استخراج ویژگی‌های ROI[3] از تصاویر با Faster R-CNN ، از بردارهای تعبیه‌شده[4] متنی تولیدشده توسط CLIP برای هر کلاس نقص نیز استفاده می‌شود. این بردارهای تعبیه شده به‌عنوان لنگرهای معنایی[5] عمل کرده و هم‌ترازی بین ROI و مفهوم کلاس را فراهم می‌آورند. این رویکرد دقت شناسایی نقص‌های کوچک و مبهم را افزایش داده و امکان بازنمایی بهتر ویژگی‌های پیچیده و معنایی هر ناحیه تصویر را فراهم می‌آورد.

مزیت‌های کلیدی روش پیشنهادی به شرح زیر است:

۱) دقت در تشخیص نقص‌های کوچک و پیچیده: ترکیب ویژگی‌های ROI از Faster R-CNN و بردارهای تعبیه‌شده معنایی CLIP، دقت تفکیک کلاس‌ها و شناسایی نقص‌های مبهم را بهبود می‌بخشد. ۲) هم‌ترازی معنایی میان تصویر و متن: استفاده از جملات راهنما با ساختار ساده و بردارهای معنایی از پیش آموزش‌دیده، ارتباطی مستقیم و معنادار میان نواحی تصویری و مفاهیم متنی برقرار می‌سازد. این هم‌ترازی، موجب افزایش دقت و پایداری در فرایند تشخیص می‌شود . ۳) نمایش غنی‌تر ویژگی‌ها با بهره‌گیری از سازوکار توجه چندگانه: مکانیزم توجه چند سر، امکان درک و ثبت وابستگی‌های چندبعدی میان نواحی تصویری و بردارهای معنایی را فراهم می‌سازد. نتیجه آن، نمایش نهایی ویژگی‌ها به‌صورت دقیق‌تر، جامع‌تر و با قدرت تفکیک بالاتر خواهد بود. ٤) بهبود فرایند آموزش با بهره‌گیری از تابع ضرر ترکیبی: استفاده هم‌زمان از ضرر مرتبط با تشخیص ناحیه‌ای و ضرر سطح کلان تصویر، موجب همگرایی بهتر مدل در مرحله آموزش شده و دقت نهایی را افزایش می‌دهد.

روش پیشنهادی با تلفیق ویژگی‌های ناحیه‌ای، بازنمایی معنایی و توجه چند سر، چارچوبی نوین برای تشخیص نقص‌های سطح

[1] Convolutional Neural Network
[2] Multi-Head Attention
[3] Region of Interest
[4] Embedding
[5] Semantic anchors

فولاد ارائه می‌دهد. نتایج تجربی حاکی از دقت خوب این روش نسبت به روش‌های مرسوم در شناسایی نقص‌های کوچک و دشوار است. این رویکرد می‌تواند به‌عنوان ابزاری مؤثر برای خودکارسازی بازرسی‌های صنعتی و پایه‌ای برای تحقیقات آتی در زمینه ارتقای تشخیص هوشمند نقص‌های سطحی مورداستفاده قرار گیرد.

ساختار مقاله به‌صورت زیر تنظیم شده است: در بخش ۲ مروری بر روش‌های پیشین، مطالعات مرتبط و چالش‌های موجود در تشخیص نقص‌های سطحی فولاد بررسی شده‌اند. بخش ۳ روش پیشنهادی به معرفی تفصیلی چارچوب SSLA-Net و اجزای اصلی آن شامل مدل پایه تشخیص، استخراج ویژگی‌های معنایی، ماژول تلفیقی و توابع ضرر اختصاص دارد. در بخش ٤ نتایج و بحث عملکرد مدل پیشنهادی ارزیابی شده و تحلیل نتایج ارائه می‌شود. در بخش ۵ نتیجه‌گیری جمع‌بندی پژوهش ارائه شده است.

۲- مروری بر روش‌های پیشین

در سال‌های اخیر، استفاده از روش‌های نوین پردازش تصویر در تشخیص نقص‌های سطحی فولاد، پیشرفت‌های قابل‌توجهی را به همراه داشته است. معماری‌های پیشرفته باتکیه‌بر استخراج ویژگی‌های دقیق از تصاویر، توانسته‌اند عملکرد مناسبی در شناسایی و طبقه‌بندی انواع نقص‌ها ارائه دهند. بااین‌حال، چالش‌هایی نظیر دشواری در تشخیص نقص‌های ریز و پیچیده، شباهت ظاهری میان کلاس‌های مختلف، و ضعف در نمایش معنایی نواحی تصویری، همچنان مانعی برای دستیابی به‌دقت بالا در شرایط صنعتی واقعی محسوب می‌شوند. علاوه بر این، بسیاری از روش‌های موجود در حفظ تعادل میان دقت و سرعت پردازش با محدودیت‌هایی مواجه‌اند. در پاسخ به این مسائل، رویکردهای جدید با تمرکز بر طراحی معماری‌های ترکیبی و چندمرحله‌ای، تلاش کرده‌اند تا با تلفیق اطلاعات ناحیه محور، ویژگی‌های معنایی و سازوکارهای توجه، دقت و پایداری تشخیص را بهبود بخشند. روش پیشنهادی این پژوهش نیز در همین راستا، باهدف ارتقا عملکرد در محیط‌های صنعتی و غلبه بر محدودیت‌های موجود طراحی شده است.

در سال ۲۰۲٤، روشی مبتنی بر یادگیری عمیق برای تشخیص عیوب سطحی فولاد ارائه شده است [۹]. در این پژوهش، با بهره‌گیری از شبکه‌های عصبی کانولوشنی مختلف از جمله EfficientNetB3 [۱۰] و MobileNetV2 [۱۱] به مقایسه کارایی این معماری‌ها در شناسایی عیوب سطحی پرداختند. نتایج نشان داد که معماری‌های مورد استفاده توانایی بالایی در شناسایی عیوب سطحی دارند. مزیت اصلی این روش‌ها، اتوماسیون فرایند بازرسی و کاهش وابستگی به نیروی انسانی است که منجر به افزایش سرعت و دقت کنترل کیفیت در صنایع فولاد می‌شود. بااین‌حال، پژوهش مذکور اشاره می‌کند که چالش‌هایی همچون نیاز به توازن داده‌ها و هزینه‌های محاسباتی در معماری‌های عمیق همچنان وجود دارد و برای بهبود بیشتر عملکرد، استفاده از روش‌هایی مانند افزایش داده یا بهره‌گیری از معماری‌های سبک‌تر پیشنهاد می‌شود.

هان و کوتبای در سال ۲۰۲۵ روشی مبتنی بر یادگیری عمیق پیشنهاد کرده‌اند که شامل پیش‌پردازش تصاویر دیتاست NEU-DET با تکنیک‌هایی مانند عملیات مورفولوژیکی، اضافه‌کردن نویز گاوسی، تحلیل مؤلفه اصلی[1] (PCA) برای کاهش بعد، و آستانه‌گذاری اوتسو برای جداسازی نقص‌ها از زمینه، و سپس آموزش مدل MobileNetV2 بر روی دیتاست جدید با تصاویر باینری و اندازه کوچک‌تر است [۲]. نتایج نشان‌دهنده دستیابی به‌دقت متوسط ۸۷٪ در طبقه‌بندی شش کلاس نقص است. این روش با کاهش اندازه داده‌ها و کاهش هزینه‌های محاسباتی، مزیت قابل‌توجهی در کاربردهای صنعتی ارائه می‌دهد؛ بااین‌حال، وابستگی به تکنیک‌های پیش‌پردازش خاص ممکن است محدودیت‌هایی در تعمیم‌پذیری به دیتاست‌های متنوع‌تر ایجاد کند، و عملکرد آن در شرایط نوری پیچیده نیازمند ارزیابی بیشتر است.

روش HCT-Det یک مدل مبتنی بر ترکیب سلسله‌مراتبی ویژگی‌های شبکه‌های کانولوشنی و ترنسفورمر برای تشخیص نقص سطح فولاد ارائه می‌دهد [٦]. این مدل از بلوک WSA-R[2] برای به‌کارگیری توجه خودی پنجره‌ای با هدف کاهش هزینه محاسباتی، و از بلوک‌های رزیدوال موازی برای حفظ پیوستگی ویژگی‌های محلی بهره می‌گیرد. ویژگی‌های تولیدشده در سطوح

[1] Principal Component Analysis

[2] Window-based Self-Attention / Residual Structure

مختلف به‌صورت مقیاس متقاطع به انکودر منتقل می‌شوند و تعامل درون مقیاس و بین مقیاس به بهبود دقت در شناسایی اهداف کوچک و متنوع کمک می‌کند. همچنین، مکانیزم Soft IoU-Aware برای هم‌ترازی دقیق امتیازات طبقه‌بندی به‌کارگرفته‌شده است. نتایج تجربی نشان می‌دهند که این روش در مقایسه با مدل‌های پیشرفته موجود عملکرد بهتری دارد، هرچند پیچیدگی محاسباتی بالا، کاربرد آن را در شرایط زمان واقعی و تصاویر با وضوح‌بالا محدود می‌سازد.

روش RSTD-YOLOv7 با هدف ارتقا عملکرد مدل YOLOv7 [۱۲] در تشخیص نقص‌های سطحی فولاد طراحی شده است[۷].در این معماری، ماژول RFBVGG[1] و مکانیزم توجه SimAM [۱۳] در شبکه پس‌زمینه[2] به کار گرفته شده‌اند تا از افت اطلاعات بافتی جلوگیری شود و انسجام ویژگی‌های محلی حفظ گردد. همچنین، در شبکه میانی[3]، ماژول STRVGG مبتنی بر ساختار ترنسفورمر Swin [۱۴] معرفی شده است که نقش مهمی در استخراج دقیق‌تر ویژگی‌های عمیق و کاهش ازدست‌رفتن اطلاعات دارد. در ادامه، سر[4] تشخیص بهبودیافته DSDH نیز برای افزایش دقت طبقه‌بندی و تسریع همگرایی مدل مورد استفاده قرار گرفته است. این ترکیب ماژول‌ها موجب بهبود قابل‌توجه در شناسایی نقص‌های کوچک و با ویژگی‌های نامشخص شده و قابلیت کاربرد صنعتی این روش را افزایش داده است. بااین‌حال، وابستگی به اجزای خاص و پیچیدگی محاسباتی بالا، ممکن است در برخی سناریوهای زمان واقعی و تصاویر با وضوح‌بالا محدودیت‌هایی ایجاد کند و نیاز به بهینه‌سازی بیشتر برای تعمیم‌پذیری در محیط‌های صنعتی متنوع‌تر احساس می‌شود.

در سال ۲۰۲۵، پژوهشی توسط اشرفی و همکاران ارائه شد که از مدل‌های پیشرفته نظیر U-Net [۱۵]، FCN-8 [۱۶] و FPN [۱۷] بهره گرفته شد که با شبکه‌های پس‌زمینه از پیش آموزش‌دیده‌ای مانند EfficientNet-B0 [۱۰] و ResNet34 [۱۸] ترکیب شده بودند. همچنین، مدل YOLOv4 [۱۹] برای تشخیص اولیه نقص‌ها بر روی مجموعه‌داده SEVERSTAL [۲۰] به کار گرفته شد [۵]. در ادامه، ساختاری دومرحله‌ای طراحی گردید که ابتدا با استفاده از YOLOv4 محل نقص‌ها را شناسایی کرده و سپس تصاویر برش خورده حاصل از مرحله اول توسط مدل U-Net مورد تقسیم‌بندی دقیق قرار گرفتند. این ترکیب هوشمندانه از تشخیص و تفکیک، به‌ویژه در شناسایی نقص‌های کوچک و پراکنده، موجب افزایش چشمگیر دقت گردید. با وجود این مزایا، عملکرد مدل در دسته‌هایی با نمونه‌های اندک، تحت‌تأثیر عدم تعادل داده‌ها قرار گرفت؛ موضوعی که لزوم بهره‌گیری از راهکارهای بهبود تعادل کلاس‌ها را برجسته می‌سازد.

یانگ و لیو روشی بهینه‌شده مبتنی بر معماری RetinaNet [۲۱] ارائه کرده‌اند که در این روش، با بهره‌گیری از کانولوشن‌های تغییرپذیر [۲۲] در ساختار اصلی شبکه ResNet طراحی شده است که امکان تنظیم پویا و تطبیقی شکل کرنل‌های کانولوشنی را فراهم می‌سازد [۸]. این ویژگی موجب استخراج دقیق‌تر ویژگی‌های نقص‌هایی با اشکال نامنظم و متنوع می‌گردد. علاوه بر آن، ماژول CA_BiFPN [۲۳] برای تلفیق ویژگی‌های چند مقیاسی معرفی شده است که با استفاده از مکانیسم توجه، تمرکز شبکه را بر نواحی مرتبط با نقص‌ها افزایش داده و در نتیجه دقت تشخیص را بهبود می‌بخشد. همچنین، تابع زیان IA-BCELoss [۲۴] با ترکیب مؤلفه‌های طبقه‌بندی و رگرسیون، موجب ارتقا کیفیت باکس‌های مرزی و بهبود عملکرد نهایی مدل می‌شود. این روش دقت بالایی دارد، بااین‌حال، پیچیدگی محاسباتی ناشی از استفاده از کانولوشن‌های تغییرپذیر ممکن است در سناریوهای زمان واقعی با منابع محدود، چالش‌هایی ایجاد کند.

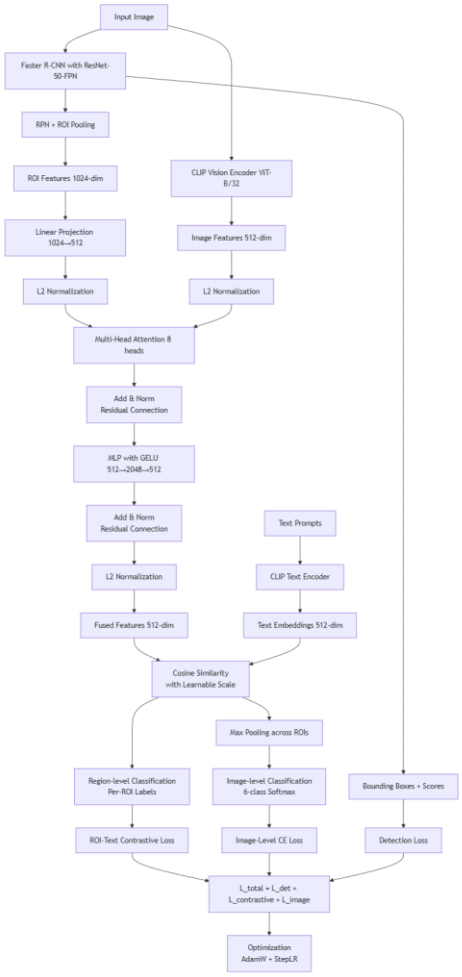
در حالت کلی بررسی مطالعات اخیر در حوزه تشخیص نقص‌های سطحی فولاد نشان‌دهنده پیشرفت چشمگیر در بهره‌گیری از الگوریتم‌های یادگیری عمیق است. مدل‌های مختلف با استفاده از تکنیک‌های تقسیم‌بندی معنایی، تشخیص شیء و مکانیسم‌های توجه، توانسته‌اند بخشی از چالش‌های موجود مانند تنوع شکل نقص‌ها، شباهت به پس‌زمینه و ابعاد کوچک آن‌ها را برطرف کنند. بااین‌حال، محدودیت‌هایی نظیر دقت پایین در تشخیص نقص‌های ریز، پیچیدگی محاسباتی در کاربردهای زمان واقعی، عدم تعادل داده‌ها و ضعف در تعمیم‌پذیری در شرایط صنعتی متنوع همچنان

[1] Receptive Field Block with VGG-style backbone
[2] Backbone
[3] Neck
[4] Head

پابرجاست. این موارد، نیاز به توسعه راهکارهای نوآورانه برای دستیابی به تعادل میان دقت، سرعت و بهره‌وری محاسباتی را برجسته می‌سازد.

۳- روش پیشنهادی

در این پژوهش، یک چارچوب ترکیبی برای تشخیص و طبقه‌بندی نقص‌های سطحی فولاد ارائه شده است که مبتنی بر تلفیق مدل‌های تشخیص ناحیه، و نمایش‌های معنایی است. مراحل روش پیشنهادی در شکل (۱) نشان‌داده‌شده است. این روش شامل مراحل زیر است:



شکل (۱): معماری روش پیشنهادی

۳-۱ مدل پایه تشخیص نقص

در این پژوهش، برای شناسایی نقص‌های سطح ورق‌های فولادی از معماری Faster R-CNN [۲۵] استفاده شده است. این معماری شامل دو بخش اصلی است: (۱) استخراج ویژگی‌ها، (۲) تولید نواحی پیشنهادی و سپس طبقه‌بندی دقیق این نواحی و تصحیح مختصات جعبه‌های محدودکننده[1]. این ساختار به افزایش دقت در شناسایی نواحی کوچک، نامنظم و بافت‌محور که در تصاویر صنعتی فولاد رایج است کمک می‌کند.

مرحله اول شامل استفاده از شبکه عمیق ResNet50 [۱۸] به‌عنوان شبکه پس‌زمینه است که پیش‌تر روی مجموعه‌داده ImageNet آموزش دیده است. ساختار این شبکه باعث استخراج ویژگی‌های غنی و چندسطحی می‌گردد که قادر به مدل‌سازی تغییرات ظریف مانند خطوط نازک و بافت‌های با وضوح کم است. لایه‌های ابتدایی این شبکه در فرایند آموزش ثابت نگه داشته شدند تا دانش عمومی حفظ شود و لایه‌های انتهایی برای تطبیق با داده‌های خاص نقص‌های سطح فولاد مجدداً آموزش داده شده‌اند که منجر به تسریع آموزش و افزایش دقت نهایی مدل می‌شود. انتخاب ResNet50 به‌عنوان شبکه پس‌زمینه کمک می‌کند تا مدل ضمن حفظ دقت بالا، توانایی استخراج ویژگی‌های پیچیده و خاص نقص‌های سطح فولاد را داشته باشد.

در مرحله دوم، شبکه پیشنهاد ناحیه[2] (RPN) با الگوریتم‌های مبتنی بر لنگر نواحی مستعد وجود نقص را تولید کرده و سپس با استفاده از روش حذف بیشینه تداخل غیر حداکثری[3](NMS) جعبه‌های تکراری حذف می‌گردند تا تنها نامزدهای مهم حفظ شوند. شبکه پیشنهاد ناحیه از نه لنگر اولیه با ترکیب سه مقیاس و سه نسبت ابعاد بهره می‌گیرد و در هر تصویر حدود ۳۰۰ ناحیه پیشنهادی تولید می‌کند

عملیات تجمیع ناحیه موردعلاقه بر روی این نواحی انجام شده و ویژگی‌های هر ناحیه استخراج می‌شود. طبقه‌بندی نهایی از طریق دولایه کاملاً متصل با ۱۰۲۴ و ۵۱۲ نورون و تابع فعال‌ساز ReLU انجام می‌شود. این رویکرد امکان بررسی دقیق و مستقل هر ناحیه

[1] Bounding box
[2] Region Proposal Network
[3] Non-Maximum Suppression

را فراهم می‌آورد و دقت قابل‌توجهی نسبت به مدل‌های تک‌مرحله‌ای مانند YOLO و SSD در شناسایی نقص‌های ریز، پراکنده و پیچیده – به‌ویژه در شرایط نویز صنعتی و تنوع بافت بالا – ارائه می‌دهد.

۳-۲ استخراج ویژگی‌های معنایی

در این پژوهش برای ارتقا درک معنایی مدل و افزایش دقت در طبقه‌بندی نقص‌های سطحی، از مدل CLIP[1] [۲۶] استفاده شده است. CLIP که توسط OpenAI معرفی شده، با بهره‌گیری از یادگیری میان تصویر و متن، قادر است نمایش‌های برداری مشترکی از داده‌های تصویری و زبانی تولید کند.

در این چارچوب، برای هر کلاس نقص سطحی فولاد، یک جمله توصیفی متنی طراحی می‌شود. به‌عنوان‌مثال، برای کلاس crazing جمله‌ای مانند "a close-up photo of crazing defect" در نظر گرفته شده است. این جملات توسط بخش متنی CLIP به بردارهای تعبیه‌شده زبانی تبدیل می‌شوند که نمایانگر معنای مفهومی هر کلاس هستند. به طور هم‌زمان، نواحی تصویری استخراج‌شده از مرحله تشخیص نیز توسط بخش تصویری CLIP پردازش شده و به بردارهای تصویری متناظر تبدیل می‌گردند.

مدل CLIP مورد استفاده در این پژوهش بر پایه معماری ترنسفورمر ViT-B/32 [۲۷] طراحی شده است. در این ساختار، تصویر ورودی به قطعات با ابعاد ۳۲×۳۲ تقسیم می‌شود و سپس با بهره‌گیری از مکانیزم توجه، نمایش‌های معنایی سطح بالا استخراج می‌گردد. برخلاف شبکه‌های کانولوشنی که عمدتاً بر ویژگی‌های محلی تمرکز دارند، معماری‌های ترنسفورمر مانند ViT قادرند ارتباطات بلند برد میان بخش‌های مختلف تصویر را مدل‌سازی کنند. قابلیتی که در شناسایی دقیق‌تر عیوب سطحی پیچیده و پراکنده نقش کلیدی ایفا می‌کند.

ویژگی منحصربه‌فرد CLIP، ایجاد یک فضای مشترک معنایی میان تصویر و متن است. در این فضا، هر ناحیه پیشنهادی از تصویر به‌صورت یک بردار تعبیه‌شده بازنمایی می‌شود و با تمام بردارهای متنی مربوط به کلاس‌ها مقایسه می‌گردد. شباهت محاسبه‌شده نشان می‌دهد که هر ناحیه به کدام کلاس متنی نزدیک‌تر است و در نهایت، برچسبی که بیشترین شباهت را دارد به آن ناحیه اختصاص داده می‌شود. به‌این‌ترتیب، به‌جای یک طبقه‌بندی صریح سنتی، تخصیص برچسب بر اساس تطابق معنایی در فضای مشترک تصویر–متن انجام می‌شود. این قابلیت موجب می‌شود مدل فراتر از ویژگی‌های صرفاً بصری عمل کرده و درک مفهومی عمیق‌تری از نقص‌ها به دست آورد.

۳-۳ طراحی ماژول تلفیقی

برای دستیابی به نمایشی ترکیبی از نقص‌های سطحی فولاد، یک ماژول تلفیقی مبتنی بر مکانیزم توجه چندسری طراحی شده است. هدف این ماژول، ادغام اطلاعات مکانی حاصل از خروجی‌های ناحیه‌ای مدل Faster R-CNN با ویژگی‌های معنایی استخراج شده از CLIP است تا مدل بتواند تصمیم‌گیری دقیق‌تری بر اساس هر دو نوع داده انجام دهد.

در این ساختار، ویژگی‌های ناحیه‌ای با ابعاد ۱۰۲۴ ابتدا از طریق یک لایه خطی به فضای برداری ۵۱۲ نگاشته شده و سپس نرمال‌سازی می‌شوند. بردارهای تعبیه شده CLIP نیز به‌صورت جداگانه نرمال‌سازی می‌شوند. مکانیزم توجه چندسری با ۸ سر و بردارهای تعبیه‌شده با ابعاد ۵۱۲، برای مدل‌سازی روابط میان نواحی تصویری و ویژگی‌های معنایی به کار گرفته می‌شود.

خروجی توجه از طریق اتصالات باقیمانده با ورودی ترکیب شده و سپس توسط یک شبکه MLP پردازش می‌شود. این شبکه شامل دو لایه کاملاً متصل است. در گام نخست، ابعاد به ۲۰۴۸ افزایش‌یافته و سپس به ۵۱۲ کاهش می‌یابد. در این فرایند از تابع فعال‌ساز GELU و نرمال‌سازی لایه‌ای استفاده شده است. خروجی نهایی نیز بار دیگر تحت نرمال‌سازی قرار گرفته و با استفاده از یک پارامتر مقیاس‌پذیر قابل‌یادگیری تنظیم می‌شود. این طراحی منجر به تولید نمایش‌های تلفیق شده‌ای می‌شود که نه‌تنها موقعیت فضایی نقص را در نظر می‌گیرند، بلکه معنای مفهومی آن را نیز بازنمایی می‌کنند. به طور مشخص، مکانیزم توجه چندسری با مدل‌سازی هم‌زمان روابط میان ویژگی‌های ناحیه‌ای و تعبیه‌های معنایی، امکان هم‌ترازسازی دقیق‌تر بین نواحی تصویری و توصیف‌های متنی را فراهم می‌سازد.

[1] Contrastive Language–Image Pretraining

نحوهٔ اتصال دقیق خروجی‌های مدل Faster R-CNN و ورودی‌های CLIP در شکل (۱) نمایش‌داده‌شده است. در این ساختار، ویژگی‌های استخراج‌شده از نواحی با ابعاد ۱۰۲٤ ابتدا از طریق یک لایه نگاشت خطی به فضای ۵۱۲-بُعدی تبدیل و نرمال‌سازی می‌شوند. بردارهای متنی حاصل از CLIP نیز در همان فضای ۵۱۲-بُعدی نرمال‌سازی شده و به ماژول توجه چندسری شامل ۸ سر توجه وارد می‌شوند. سپس خروجی توجه از طریق اتصالات باقیمانده و شبکه MLP با ساختار ۵۱۲→۲۰٤۸→۵۱۲ پردازش شده و ویژگی‌های ادغام‌شده نهایی تولید می‌گردند. بدین ترتیب، ارتباط میان نمایش‌های بصری و معنایی به‌صورت دقیق و هم‌تراز مدل‌سازی می‌شود.

در فرایند آموزش، وزن‌های مدل از پیش آموزش‌دیده CLIP در هر دو بخش تصویری و متنی ثابت نگه داشته شده‌اند تا از بروز ناپایداری گرادیان و انتقال بیش از حد جلوگیری شود. تنها پارامترهای لایه فرافکنی خطی، ماژول توجه چندسری و شبکه MLPبه‌صورت قابل‌یادگیری تنظیم گردیده‌اند. ماژول توجه چندسری با ۸ سر و ابعاد برداری ۵۱۲ پیاده‌سازی شده است و شبکه MLP شامل دو لایه کاملاً متصل با ابعاد ۵۱۲→۲۰٤۸→۵۱۲ است. برای به‌روزرسانی پارامترها از بهینه‌ساز AdamW و کاهنده نرخ یادگیری StepLR استفاده شده است. همچنین، وزن‌های ترکیبی توابع زیان به‌صورت تجربی تعیین شده‌اند تا توازن مناسبی میان جنبه‌های مکانی، معنایی و کلی تصویر برقرار گردد.

٤-۳ طراحی توابع ضرر ترکیبی

برای آموزش دقیق معماری پیشنهادی، سه تابع ضرر طراحی شده‌اند که به‌صورت هم‌زمان جنبه‌های مکانی، معنایی و کلی تصویر را پوشش می‌دهند. نخست، تابع استاندارد مدل Faster R-CNN برای تشخیص و موقعیت‌یابی نقص‌ها حفظ شده است. سپس، یک تابع میان ویژگی‌های ناحیه‌ای و بردارهای تعبیه‌شده متنی CLIP تعریف شده تا مدل بتواند نواحی تصویری را با توصیف‌های معنایی هم‌راستا کند. در نهایت، از یک تابع آنتروپی متقاطع در سطح تصویر استفاده شده تا مدل کلاس غالب نقص را در کل تصویر تشخیص دهد.

ترکیب وزن‌دار این توابع، مدل را به سمت یادگیری چندوجهی هدایت می‌کند و موجب افزایش دقت، کاهش ابهام در طبقه‌بندی، و بهبود عملکرد در شرایط صنعتی می‌شود. توابع ضرر به‌صورت زیر هستند:

تابع ضرر تشخیص شیء:

$$\ell_{\det} = \ell_{cls} + \lambda_1 \ell_{reg} \quad (۱)$$

$$\ell_{cls} = -\sum_i y_i \log(p_i)$$

$$\ell_{reg} = -\sum_i Smooth_{L1}(t_i - t_i^*)$$

تابع ضرر ناحیه – متن:

$$\ell_{contrast} = -\log \frac{\exp(sim(v_i, t_i)/\tau)}{\sum_j \exp(sim(v_i, t_j)/\tau)} \quad (۲)$$

تابع ضرر سطح تصویر:

$$\ell_{img} = -\sum_k y_k \log(p_k) \quad (۳)$$

تابع ضرر نهایی ترکیبی:

$$\ell_{total} = \alpha \ell_{\det} + \beta \ell_{contrast} + \gamma \ell_{img} \quad (٤)$$

در نهایت روش پیشنهادی با تلفیق مدل‌های تشخیص ناحیه، نمایش‌های معنایی و ماژول توجه، چارچوبی چندلایه برای شناسایی دقیق نقص‌های سطحی فولاد ارائه می‌دهد. طراحی توابع ضرر ترکیبی، یادگیری هم‌زمان ویژگی‌های مکانی، مفهومی و سطح تصویر را ممکن ساخته است. این ساختار، پایه‌ای منسجم برای ارزیابی عملکرد مدل در شرایط صنعتی واقعی فراهم می‌کند.

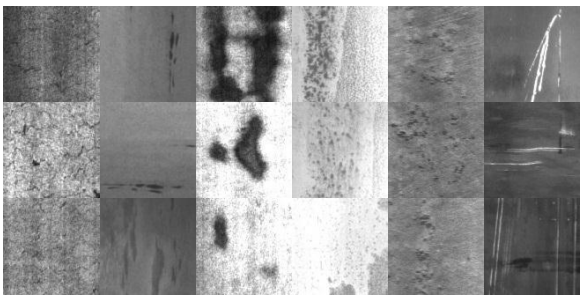
٤- نتایج و بحث

در این بخش، عملکرد معماری پیشنهادی در شناسایی نقص‌های سطحی فولاد مورد ارزیابی قرار گرفته است. ارزیابی‌ها بر اساس مجموعه‌داده صنعتی NEU-DET [۳] انجام شده و شامل تحلیل نتایج در دو سطح تشخیص ناحیه‌ای و طبقه‌بندی سطح تصویر است.

٤-۱ معرفی مجموعه‌داده

برای ارزیابی عملکرد معماری پیشنهادی، از مجموعه‌داده استاندارد NEU-DET استفاده شده است. این مجموعه‌داده شامل تصاویر خاکستری از سطح فولاد بوده و به طور خاص برای تشخیص نقص‌های سطحی صنعتی طراحی شده است. NEU-DET شامل

۶ کلاس نقص رایج در فولاد صنعتی است. این کلاس‌ها شامل پوسته‌های نورد شده[1] (RS)، لکه‌ها[2](Pa) ، ترک‌های سطحی[3] (Cr)، سطح حفره‌دار[4](PS) ، ناخالصی‌ها[5](In) و خراش‌ها[6] (Sc) هستند. این مجموعه‌داده در مجموع شامل ۱،۸۰۰ تصویر است که به‌صورت پیش‌فرض به ۱،۴۴۰ تصویر برای آموزش (Train) و ۳۶۰ تصویر برای تست (Test) تقسیم‌بندی شده‌اند. این تقسیم‌بندی توسط سازندگان مجموعه‌داده انجام شده و تنوع بالایی در ظاهر نقص‌ها را پوشش می‌دهد. در شکل (۲)، نمونه‌هایی از هر کلاس نشان‌داده‌شده است.

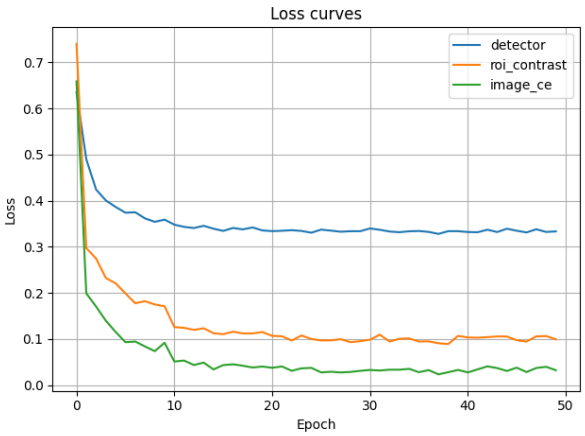


شکل (۲): تعدادی از تصاویر دیتاست NEU-DET در شش کلاس مختلف

۴-۲ آماده‌سازی داده‌ها، پیش‌پردازش و آموزش

برای افزایش تنوع داده‌ها و ارتقای توانایی مدل در تعمیم به شرایط واقعی، مجموعه‌ای از تبدیلات تصویری بر روی داده‌ها اعمال شده است. این تبدیلات شامل چرخش‌های تصادفی، تغییرات در روشنایی و کنتراست، اعمال نویز گوسی، و برش‌های موضعی بوده‌اند. هدف از این فرایند، شبیه‌سازی شرایط متغیر محیط صنعتی و افزایش مقاومت مدل در برابر تغییرات ظاهری بوده است، بدون آن که ساختار نقص‌ها دچار اختلال شود. روش پیشنهادی در ۵۰ ایپوک آموزش داده شده است. فرایند آموزش با استفاده از **بهینه‌ساز** AdamW انجام شد که برای بخش Faster R-CNN نرخ یادگیری برابر با 1e-5 و برای بخش ماژول تلفیقی نرخ یادگیری برابر با 1e-4 در نظر گرفته شد. مقدار ضرایب توابع ضرر یکسان در نظر گرفته شده است. مقدار اندازه دسته در مرحله آموزش برابر با ۴ و در مرحله اعتبارسنجی برابر با ۲ تنظیم شد .

برای بررسی روند همگرایی مدل پیشنهادی، نمودار تغییرات سه تابع ضرر اصلی در طول مراحل آموزش رسم شده است. همان‌طور که در شکل (۳) مشاهده می‌شود، هر سه تابع ضرر در طول آموزش کاهش یکنواخت و پایداری داشته‌اند. تابع ضرر تشخیص ناحیه‌ای که با رنگ آبی نمایش داده شده، از مقدار اولیه حدود ۰/۶۳ به حدود ۰/۳۳ همگرا شده است، تابع تطابق ناحیه–متن با رنگ نارنجی از مقدار اولیه ۰/۷ به حدود ۰/۰۹ کاهش یافته، و تابع طبقه‌بندی سطح تصویر با رنگ سبز از ۰/۶۵ به حدود ۰/۰۳ رسیده است. این روند نشان‌دهنده‌ی بهینه‌سازی مؤثر هر بخش از مدل و تعامل مناسب میان وظایف چندگانه در معماری پیشنهادی است، به‌گونه‌ای که کاهش هم‌زمان و پایدار هر سه تابع ضرر بیانگر یادگیری هماهنگ و هم‌افزای مدل در سطوح مختلف تشخیص و طبقه‌بندی است.



شکل (۳): نمودار تغییرات سه تابع ضرر اصلی مدل پیشنهادی

۴-۳ نتایج تشخیص ناحیه‌ای

برای ارزیابی عملکرد مدل در سطح تشخیص ناحیه‌ای، از معیار میانگین دقت متوسط در آستانه ۵۰ درصد هم‌پوشانی (mAP@50) استفاده شده است. این معیار نشان می‌دهد که مدل تا چه حد قادر است نواحی دارای نقص را به‌درستی شناسایی و طبقه‌بندی کند. این امر با درنظرگرفتن تطابق هندسی میان جعبه‌های

[1] Rolled-in Scale
[2] Patches
[3] Crazing
[4] Pitted Surface
[5] Inclusion
[6] Scratches

محدودکننده پیش‌بینی‌شده و جعبه‌های محدودکننده واقعی صورت می‌گیرد. این معیار با استفاده از معادله (۵) محاسبه می‌شود:

$$AP_c = \int_0^1 P_c(r)dr \qquad (۵)$$

$$mAP_{50} = \frac{1}{C}\sum_{c=1}^{C} AP_c$$

$$IoU = \frac{|B_{pred} \cap B_{gt}|}{|B_{pred} \cup B_{gt}|}$$

نتایج حاصل از ارزیابی روش پیشنهادی در مقایسه با چندین روش مطرح در جدول (۱) ارائه شده است. در این جدول، مقادیر mAP@50 برای هر یک از کلاس‌های شش‌گانه نقص و همچنین میانگین کلی آن گزارش شده است. مقدار mAP@50 کلی روش پیشنهادی ٪۷۸/٤ به دست آمده که آن را در جایگاه دومین روش برتر قرار می‌دهد. نکته قابل‌توجه فاصله کمتر از ٪۱ (٪۰/۹) با روش برتر RSTD-YOLOv7 با مقدار ٪۷۹/۳ است که نشان‌دهنده عملکرد بسیار رقابتی روش پیشنهادی است. علاوه بر معیار mAP@50، به‌منظور ارزیابی دقت مدل در همپوشانی‌های سخت‌گیرانه‌تر، مقدار mAP@75 نیز محاسبه شد که مقدار ٪٤۸/۰۳ به دست آمد. این نتیجه نشان می‌دهد مدل توانایی حفظ دقت مکانی را حتی در شرایط همپوشانی بالا دارد که برای کاربردهای صنعتی اهمیت ویژه‌ای دارد.

برخلاف برخی روش‌ها که تنها در شناسایی دسته‌های خاص موفق عمل می‌کنند، روش پیشنهادی توازن دقت قابل‌قبولی از خود نشان می‌دهد. این سطح از پایداری و تعادل در دقت، پتانسیل بالای روش پیشنهادی را برای کاربرد در محیط‌های صنعتی واقعی – که تنوع و پیچیدگی نقص‌ها یک چالش اساسی است – به‌خوبی نشان می‌دهد.

جدول (۱): نتایج حاصل از تشخیص ناحیه‌ای در روش‌های متفاوت

روش	cr	in	pa	ps	rs	sc	mAP (%)
[٤]	۷/٤۱	۷٦/۳	۸٦/۳	۸۵/۱	۵۸/۱	۸۵/٦	۷۲/٤
[۲۸]	۸۷/۵	۷۵/٤	۷٦/۲	۸۱/۷	٦۵/۳	۷۲/۲	۷٦/٤
LAACNet [۲۹]	٤۷/۱	۸۸/٤	۹۲/۸	۷٦/۵	۵۹/۳	۹۲/۲	۷٦/۰
[۷RSTD-YOLOv7]	۵۱/۲	۷۸/۱	۹۲/۷	۹۲/٤	٦۷/۲	۹٤/۰	۷۹/۳
[۳۰]	۵۹/۱	۸٦/۷	۹۵/۱	۸۵/۹	٦۵/٦	۷۲/٦	۷۷/۵
[۳۲]	۵۵/۱	۸۳/۰	۹۳/٦	۸٦/۱	۵۹/۷	۸٤/۲	۷۷/۰
[۳۳]	۵۱/۸	۸۱/۵	۹۰/۵	۹۰/۰	۵۹/٤	۸٦/۵	۷٦/٦
[۳٤]	-	-	-	-	-	-	۷۳/۸
[۳٦]	٦۵/۲	۸۲/۰	۸۷/۲	۷٤/۰	۷٦/۸	۷٤/٤	۷٦/٦
[۳۸]	-	-	-	-	-	-	۷۸/۲
روش پیشنهادی	۵۸/۹	۸۲/٤	۹۳/۸	۸۵/۲	۵۹/۳	۹۰/۳	۷۸/٤

برای ارزیابی کیفی عملکرد مدل، مجموعه‌ای از تصاویر خروجی شامل جعبه‌های محدودکننده واقعی و پیش‌بینی‌شده در شکل (٤) نشان‌داده‌شده است. در این تصاویر، جعبه‌های قرمز نشان‌دهنده موقعیت واقعی نقص‌ها بوده و جعبه‌های سبز موقعیت‌های پیش‌بینی‌شده توسط مدل را نمایش می‌دهند. عددی که روی هر جعبه پیش‌بینی‌شده نوشته شده است، درجه اعتماد مدل به پیش‌بینی آن ناحیه است که میزان اطمینان مدل نسبت به تعلق ناحیه به یک کلاس خاص را نشان می‌دهد. این مقدار از خروجی لایه طبقه‌بندی مدل استخراج شده و با فرمول زیر محاسبه می‌شود:

$$confidence = \max_i P(class_i \mid box) \qquad (٦)$$

مطابق با شکل (٤) در تصویر اول، مدل با درجه اعتماد بالا (۰/۹٤) موفق به تشخیص دقیق یک ناحیه در کلاس in شد که تطابق هندسی مطلوبی با داده واقعی داشت. تصویر دوم شامل سه ناحیه pa بود که دو مورد با درجه اعتماد بالا (۰/۹۷ و ۰/۸۵) به‌درستی شناسایی شدند، درحالی‌که مورد سوم با درجه اعتماد پایین‌تر (۰/٦۱) نشان‌دهنده‌ی ابهام یا ضعف در تشخیص است. در تصویر

سوم، مدل چهار ناحیه sc را شناسایی کرده است که سه مورد با درجه اعتماد بسیار بالا (تا ۰/۹۹) تطابق خوبی با جعبه‌های واقعی داشتند، اما یک مورد با درجه اعتماد پایین (۰/۳۷) یک پیش‌بینی اشتباه است. این تصاویر نشان می‌دهند که مدل در تشخیص دقیق و طبقه‌بندی نواحی دارای نقص عملکردی قابل‌اتکا دارد، هرچند تنظیم آستانه می‌تواند به کاهش خطاهای جزئی کمک کند.

٤-٤ نتایج طبقه‌بندی تصویر

در این بخش، عملکرد مدل پیشنهادی در طبقه‌بندی انواع نقص‌های سطحی فولاد با معیارهای معتبر و استاندارد در حوزه تشخیص خودکار عیوب صنعتی مورد بررسی و مقایسه قرار گرفته است. معیارهای متداول ارزیابی شامل دقت (Accuracy)، صحت (Precision)، بازیابی (Recall) و معیار ترکیبی F1-score است که همه به‌صورت جامع در ارزیابی عملکرد مدل‌ها لحاظ گردیده‌اند. نتایج مقایسه‌ای استخراج شده از مقالات مرجع و معتبر در این حوزه، در قالب جدول ۲ ارائه شده‌اند تا امکان تحلیل و بررسی مقاومتی میان روش‌ها فراهم شود.

بر اساس نتایج ارائه‌شده در جدول، مدل پیشنهادی در تمامی معیارهای ارزیابی عملکرد به مقدار کامل (٪۱۰۰) دست یافته است. این امر گویای توانمندی بالای مدل در شناسایی دقیق و کامل شش نوع نقص سطحی استاندارد موجود در مجموعه‌داده معروف NEU-DET بوده و نشان از همگامی کامل پیش‌بینی‌های مدل با داده‌های واقعی دارد. دومین روش برتر، روش ارائه‌شده در مرجع [۳۵] است که دقت و عملکرد بسیار نزدیک به مدل پیشنهادی دارد. روش‌های قدیمی‌تر علی‌رغم پیشرفت‌های قبلی، عملکرد مناسبی نداشته و در معیارهای ذکر شده فاصله قابل‌ملاحظه‌ای با مدل‌های نوین دارند. ضروری است ذکر شود که برخی از روش‌های مقایسه‌ای تنها دقت کلی را گزارش کرده‌اند، بدون ارائه جزئیات کامل معیارهای طبقه‌بندی که این امر مقایسه دقیق و عمیق را دشوار ساخته است.

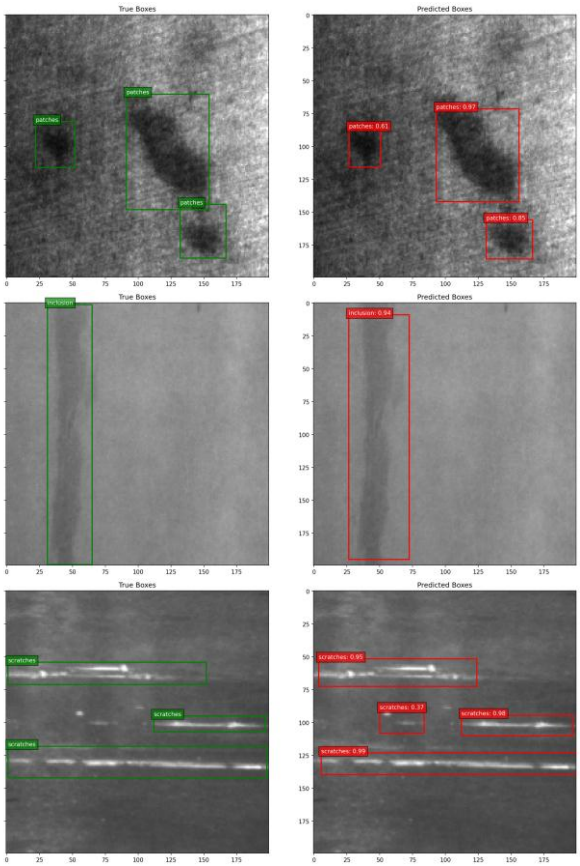
دستیابی به دقت ٪۱۰۰ در طبقه‌بندی سطح تصویر را می‌توان به سه عامل کلیدی نسبت داد:

(۱) تقویت قابلیت تفکیک مفهومی کلاس‌ها:

تعبیه‌سازی‌های معنایی CLIP تفاوت‌های مفهومی میان دسته‌های عیوب فولاد را به شکل مؤثری برجسته کرده و موجب افزایش فاصله کلاسی در فضای ویژگی شده‌اند. (۲) بهبود تشخیص نواحی کوچک و بافت‌محور:

ماژول توجه چندسر با مدل‌سازی روابط متقابل میان ROIها و متن، حساسیت مدل به جزئیات ناحیه‌ای را افزایش داده است. (۳) هم‌افزایی میان اطلاعات ناحیه‌ای و تصویر کامل:

استفاده از شباهت کسینوسی در سطح ROI و تجمیع پایدار آن در سطح تصویر، ریسک خطاهای محلی را کاهش داده و تصمیم نهایی را دقیق‌تر کرده است. این نتایج نشان می‌دهد SSLA-Net توانسته است چالش‌هایی مانند نقص‌های کوچک، بافت‌های شبیه به هم و تغییرات نوری را به‌صورت کارآمد مدیریت کند.



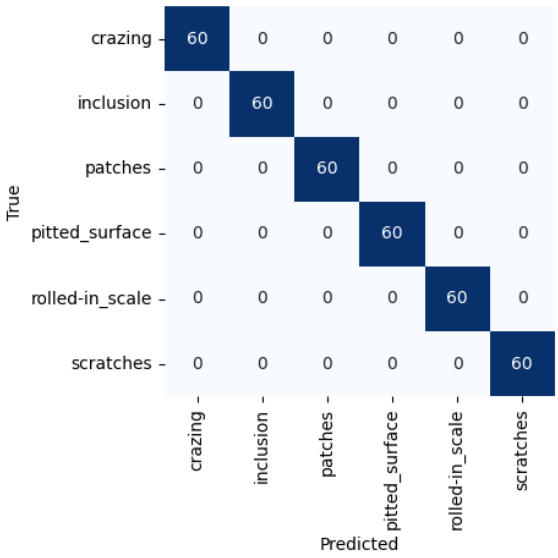
شکل (٤): مقایسه جعبه‌های محدودکننده واقعی (سبز) و پیش‌بینی‌شده (قرمز) در تشخیص نقص سطحی

ماتریس برخورد مدل پیشنهادی که در شکل (۵) نشان‌داده‌شده، توزیع پیش‌بینی‌های مدل را در هر کلاس نقص نمایش می‌دهد. در

این ماتریس، سطرها نمایانگر کلاس‌های واقعی و ستون‌ها نمایانگر کلاس‌های پیش‌بینی شده هستند. نکته برجسته در ماتریس مدل پیشنهادی، صفر بودن کلیه مقادیر خارج از قطر اصلی است؛ این موضوع ضرورتاً بیانگر عدم بروز هیچ خطای طبقه‌بندی و پیش‌بینی دقیق و کامل تمامی نمونه‌ها است. اگرچه SSLA-Net عملکرد قدرتمندی نشان داده است؛ اما پیاده‌سازی ماژول توجه چندسر و پردازش ویژگی‌های ROI هزینه محاسباتی بیشتری نسبت به تشخیص‌گرهای تک‌مرحله‌ای دارد.

جدول (۲): نتایج حاصل از طبقه‌بندی تصویر در روش‌های متفاوت

روش	Precision	Recall	F1-score	Acc
[۲]	۰/۸۸	۰/۸۷	۰/۸۷	۰/۸۷
[۳۵]	۹۹/۳۳	۹۹/۳۳	۹۹/۱۷	۹۹/۴۴
[۳۵]	–	–	–	۸۹/۶۰
[۳۹]	–	–	–	۹۹/۰۵
[۴۰]	–	–	–	۹۵/۰۰
[۴۱]	۹۷/۸۱	۹۷/۷۷	۹۷/۷۸	۹۷/۷۷
[۴۲]	-	-	-	۹۷/۲۲
InceptionResNet-V2 [۴۳]	۱۰۰	۱۰۰	۱۰۰	۱۰۰
روش پیشنهادی	۱۰۰	۱۰۰	۱۰۰	۱۰۰



شکل (۵): ماتریس برخورد مربوط به روش پیشنهادی

۵- نتیجه‌گیری

در این پژوهش، یک چارچوب ترکیبی برای تشخیص و طبقه‌بندی نقص‌های سطحی فولاد طراحی و ارزیابی شده است که با تلفیق معماری Faster R-CNN، نمایش‌های معنایی مدل CLIP و ماژول توجه چندسری، عملکردی دقیق و چندلایه در شرایط صنعتی واقعی ارائه می‌دهد. استفاده هم‌زمان از اطلاعات مکانی و مفهومی، موجب افزایش دقت در شناسایی نواحی کوچک، پراکنده و بافت‌محور شده و امکان طبقه‌بندی معنایی دقیق‌تر را فراهم ساخته است. طراحی توابع ضرر ترکیبی نیز نقش مؤثری در هدایت مدل به سمت یادگیری هم‌زمان ویژگی‌های ناحیه‌ای، معنایی و کلی تصویر ایفا کرده است.

نتایج تجربی نشان داد که روش پیشنهادی در تشخیص ناحیه‌ای با دستیابی به مقدار mAP@50 برابر با ٪۷۸/٤ عملکردی رقابتی و متوازن نسبت به روش‌های موجود دارد. همچنین، در طبقه‌بندی سطح تصویر، مدل موفق به دستیابی به دقت ٪۱۰۰ شده است که بیانگر توانمندی آن در تشخیص کلاس غالب نقص در شرایط عملیاتی است. دقت بالای مدل به‌ویژه در طبقه‌بندی سطح تصویر، ناشی از هم‌افزایی میان نمایش‌های معنایی CLIP و ویژگی‌های مکانی استخراج‌شده از Faster R-CNN است. این ترکیب موجب افزایش تفکیک‌پذیری کلاس‌ها، بهبود تشخیص نواحی کوچک و کاهش خطاهای موضعی در تصمیم‌گیری نهایی شده است. این نتایج، همراه با ساختار چندلایه و قابلیت تعمیم‌پذیری مدل، نشان‌دهنده‌ی پتانسیل بالای روش پیشنهادی برای به‌کارگیری در سامانه‌های نظارت صنعتی و توسعه‌های آتی در حوزه تشخیص نقص‌های سطحی است.

درعین‌حال، لازم به ذکر است که مدل پیشنهادی دارای برخی محدودیت‌ها نیز هست. به دلیل استفاده از معماری‌های عمیق مانند CLIP و ماژول توجه چندسری، هزینه محاسباتی مدل نسبتاً بالا است و اجرای بلادرنگ آن در محیط‌های صنعتی نیازمند سخت‌افزارهای گرافیکی پیشرفته است. به‌عنوان مسیرهای آتی، می‌توان به بهینه‌سازی شبکه برای کاهش پیچیدگی محاسباتی و افزایش سرعت اجرا، و ارزیابی عملکرد مدل بر روی داده‌های صنعتی واقعی با شرایط نوری و بافتی متنوع‌تر اشاره کرد. این اقدامات می‌توانند موجب ارتقای کاربردپذیری و قابلیت تعمیم مدل در خطوط تولید واقعی شوند.

References

[1] X. Wen, J. Shan, Y. He, and K. Song, “Steel Surface Defect Recognition: A Survey,” Coatings, vol. 13, no. 1, art. no. 17, 2022, doi: 10.3390/coatings13010017.

[2] O. C. Han and U. Kutbay, “Detection of Defects on Metal Surfaces Based on Deep Learning,” Applied Sciences, vol. 15, no. 3, art. no. 1406, 2025, doi: 10.3390/app15031406.

[3] Y. He, K. Song, Q. Meng, and Y. Yan, “An end-to-end steel surface defect detection approach via fusing multiple hierarchical features,” IEEE Trans.
Instrum. Meas., vol. 69, no. 4, pp. 1493–1504, Apr. 2020, doi: 10.1109/TIM.2019.2913672.

[4] X. Lv, F. Duan, J. J. Jiang, X. Fu, and L. Gan, “Deep metallic surface defect detection: The new benchmark and detection network,” Sensors, vol. 20, no. 6, p. 1562, Mar. 2020, doi: 10.3390/s20061562.

[5] S. Ashrafi, M. Rahim, H. Li, and A. Zhang, “Steel surface defect detection and segmentation using deep neural networks,” Results in Engineering, vol. 25, art. no. 103972, 2025, doi: 10.1016/j.rineng.2025.103972.

[6] X. Chen, L. Wang, Y. Zhao, and J. Liu, “HCT-Det: A High-Accuracy End-to-End Model for Steel Defect Detection Based on Hierarchical CNN–Transformer Features,” Sensors, vol. 25, no. 5, art. no. 1333, 2025, doi: 10.3390/s25051333.

[7] H. Song, “RSTD-YOLOv7: A steel surface defect detection based on improved YOLOv7,” Scientific Reports, vol. 15, art. no. 19649, 2025, doi: 10.1038/s41598-025-04811-w.

[8] Z. Yang and Y. Liu, “A steel surface defect detection method based on improved RetinaNet,” Scientific Reports, vol. 15, art. no. 6045, 2025, doi: 10.1038/s41598-025-88527-x.

[9] I. U. Khan, N. Aslam, M. Aboulnour, A. Bashamakh, F. Alghool, N. Alsuwayan, R. Alturaif, H. Gull, S. Z. Iqbal, and T. Hussain, "Deep Learning-Based Surface Defect Detection in Steel Products Using Convolutional Neural Networks ", Mathematical Modelling of Engineering Problems, vol. 11, no. 11, pp. 3006–3014, Nov. 2024, doi: 10.18280/mmep.111113.

[10] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” arXiv preprint arXiv:1905.11946, 2019.

[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 4510–4520.

[12] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Vancouver, Canada, Jun. 2023, pp. 7464–7475. doi: 10.1109/CVPR52729.2023.00721.

[13] L. Yang, R.-Y. Zhang, L. Li, and X. Xie, “SimAM: A simple, parameter-free attention module for convolutional neural networks,” in Proc. 38th Int. Conf. Mach. Learn. (ICML), Jul. 2021, pp. 11863–11874. [Online]. Available: https://doi.org/10.48550/arXiv.2103.06264

[14] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, “Swin Transformer: Hierarchical vision transformer using shifted windows,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2021, pp. 10012–10022. doi: 10.1109/ICCV48922.2021.00986

[15] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, Springer, 2015, pp. 234–241. [Online]. Available: https://doi.org/10.1007/978-3-319-24574-4_28

[16] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 3431–3440. [Online]. Available: https://doi.org/10.48550/arXiv.1411.4038

[17] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017, pp. 2117–2125. [Online]. Available: https://doi.org/10.48550/arXiv.1612.03144

[18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in Proc. CVPR, 2016, pp. 770–778. [Online]. Available: https://arxiv.org/abs/1512.03385

[19] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv preprint arXiv:2004.10934, 2020. [Online]. Available: https://doi.org/10.48550/arXiv.2004.10934

[20] A. Grishin, B. Bardintsev, and Severstal, “Severstal: Steel Defect Detection,” Kaggle Competition Dataset, 2019. [Online]. Available: https://www.kaggle.com/c/severstal-steel-defect-detection

[21] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2980–2988. [Online]. Available: https://doi.org/10.48550/arXiv.1708.02002

[22] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable Convolutional Networks," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 764–773. [Online]. Available: https://doi.org/10.48550/arXiv.1703.06211

[23] Q. Hou, D. Zhou, and J. Feng, "Coordinate Attention for Efficient Mobile Network Design," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2021, pp. 13713–13722. [Online]. Available: https://doi.org/10.48550/arXiv.2103.02907

[24] Q. Li, X. Jia, J. Zhou, L. Shen, and J. Duan, "Rediscovering BCE Loss for Uniform Classification," arXiv preprint arXiv:2403.07289, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2403.07289

[25] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, Jun. 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2577031

[26] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, et al., "Learning Transferable Visual Models From Natural Language Supervision," in Proc. Int. Conf. Mach. Learn. (ICML), Jul. 2021, pp. 8748–8763. [Online]. Available: https://arxiv.org/abs/2103.00020

[27] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in Proc. Int. Conf. Learn. Represent. (ICLR), May 2021. [Online]. Available: https://arxiv.org/abs/2010.11929

[28] C. Zhang, X. Bai, and X. Zhang, "Metal plate surface defect detection method based on attention mechanism," Comput. Integr. Manuf. Syst., vol. 29, no. 3, pp. xx–xx, 2023.

[29] Z. L. Lv et al., "Lightweight adaptive activation convolution network based defect detection on polished metal surfaces," Engineering Applications of Artificial Intelligence, vol. 133, p. 108482, 2024. doi: 10.1016/j.engappai.2024.108482

[30] M. Siddiqui, Y. Muhammad, and H. Ahn, "Faster metallic surface defect detection using deep learning with channel shuffling," arXiv preprint, arXiv:2406.14582, 2024.

[31] B. Li et al., "NGD-YOLO: An improved real-time steel surface defect detection algorithm," Electronics, vol. 14, no. 14, p. 2859, 2025.

[32] C. Li et al., "Steel surface defect detection method based on improved YOLOX," IEEE Access, vol. 12, pp. 37643–37652, 2024.

[33] S. Zhou, Y. Cai, Z. Zhang, and J. Yin, "MESC-DETR: An improved RT-DETR algorithm for steel surface defect detection", Electronics, vol. 14, no. 11, p. 2232, 2025.

[34] Z. Li, X. Wei, M. Hassaballah, Y. Li, and X. Jiang, "A deep learning model for steel surface defect detection", Complex & Intelligent Systems, vol. 10, no. 1, pp. 885–897, 2024.

[35] A. A. M. Ibrahim and J. R. Tapamo, "Transfer learning-based approach using new convolutional neural network classifier for steel surface defects classification," Scientific African, vol. 23, p. e02066, 2024.

[36] T. Wi, M. Yang, S. Park, and J. Jeong, "D²-SPDM: Faster R-CNN-based defect detection and surface pixel defect mapping with label enhancement in steel manufacturing processes", *Applied Sciences*, vol. 14, no. 21, p. 9836, 2024.

[37] S. Tian, P. Huang, H. Ma, J. Wang, X. Zhou, S. Zhang, J. Zhou, R. Huang, and Y. Li, "CASDD: Automatic surface defect detection using a complementary adversarial network," *IEEE Sensors Journal*, vol. 22, no. 20, pp. 19583–19595, Oct. 2022. doi: 10.1109/JSEN.2022.3202179

[38] C. Chen, H. Lee, and M. Chen, "Steel surface defect detection method based on improved YOLOv9," Scientific Reports, vol. 15, no. 1, p. 25098, 2025. DOI: 10.1038/s41598-025-10647-1

[39] L. Yi, G. Li, and M. Jiang, "An end-to-end steel strip surface defects recognition system based on convolutional neural networks," *Steel Research International*, vol. 88, no. 2, p. 1600068, Feb. 2017. doi: 10.1002/srin.201600068

[40] A. A. M. Ibrahim and J. R. Tapamo, "Steel surface defect detection and classification using bag of visual words with BRISK," in *Proc. Congress on Smart Computing Technologies*, Singapore, Dec. 2022, pp. 235–246. doi: 10.1007/978-981-99-2468-4_18

[41] M. Jeong, M. Yang, and J. Jeong, "Hybrid-DC: A hybrid framework using ResNet-50 and Vision Transformer for steel surface defect classification in the rolling process," Electronics, vol. 13, no. 22, p. 4467, Nov. 2024. DOI: 10.3390/electronics13224467.

[42] S. T. K. Ayon, F. M. Siraj, and J. Uddin, “Steel Surface Defect Detection Using Learnable Memory Vision Transformer,” *Computers, Materials & Continua*, vol. 82, no. 1, 2025.

[43] A. Bouguettaya and H. Zarzour, “CNN-based hot-rolled steel strip surface defects classification: a comparative study between different pre-trained CNN models”, Int. J. Adv. Manuf. Technol., vol. 132, no. 1, pp. 399–419, 2024.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 125-139

SSLA-Net: Semantic–Spatial Learning Architecture for Robust Steel Surface Defect Detection

Masoumeh Rezaei[1*], Mansoureh Rezaei [2], Mehri Rajaei[3]

[1]Assistant Professor, Department of Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran
[2]PhD in Artificial Intelligence, Computer Engineering Department, Yazd University, Yazd, Iran
[3] Assistant Professor, Department of Information Technology Engineering, University of Sistan and Baluchestan, Zahedan, Iran

Article Information

Original Research Paper

Received:
2025 September 13

Accepted:
2025 November 03

Keywords:
Multimodal Learning, Surface Defect Detection,Multi-Head Attention, Semantic–Spatial Representation, Industrial Vision

Corresponding Author*:
mrezaei@ece.usb.ac.ir

Abstract

In the steel industry, accurate and rapid detection of surface defects plays a vital role in ensuring product quality and reducing waste. However, the visual similarity between texture patterns and the complexity of defective regions limit the effectiveness of conventional machine vision methods. In this study, a multimodal hybrid framework is proposed for the detection and classification of steel surface defects. The proposed model integrates the region-based detection capability of Faster R-CNN with the semantic representations of CLIP, connected through a multi-head attention module that jointly models the relationship between visual and conceptual features. This synergy enhances the accuracy in detecting small and scattered defects and improves class separability in the feature space. The architecture employs three complementary loss functions — region detection loss, region–text alignment loss, and image-level loss — to enable simultaneous learning of spatial and semantic information. Experimental results on the NEU-DET benchmark dataset demonstrate that the proposed model achieves 100% accuracy in image-level classification and 78.04% mAP@50 in region-level detection. These promising results indicate that CLIP-based multimodal learning can significantly enhance the accuracy and robustness of defect detection in real industrial environments.

: 10.22034/ABMIR.2025.23624.1166

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۵۸-۱۴۱

هوشمند سازی کنترل کیفیت و تشخیص خرابی در فرآیندهای تولید فولاد با استفاده از یادگیری عمیق و پردازش تصویر به‌منظور بهینه‌سازی مصرف انرژی

فاطمه سعادت جو[1]*، مسلم ملایی [2]

[1]استادیار، گروه مهندسی کامپیوتر، دانشگاه علم و هنر، یزد، ایران

[2]معلم، آموزش و پرورش هرمزگان، گروه کامپیوتر آموزش و پرورش رودان، هرمزگان، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۶/۲۴

تاریخ پذیرش:
۱۴۰۴/۰۸/۲۶

کلیدواژه‌ها:
هوش مصنوعی، یادگیری عمیق، CNN-LSTM، صنعت فولاد

نویسنده‌مسئول:
saadatjou@sau.ac.ir

چکیده

در سال‌های اخیر، افزایش رقابت در صنعت فولاد و ضرورت کاهش هزینه‌ها در کنار بهبود کیفیت، استفاده از فناوری‌های هوشمند مبتنی بر یادگیری عمیق را بیش‌ازپیش ضروری کرده است. یکی از چالش‌های اصلی در این حوزه، کنترل کیفیت محصولات و شناسایی سریع نواقص سطحی است که مستقیماً بر بهره‌وری خطوط تولید و میزان مصرف انرژی اثر می‌گذارد. در این پژوهش یک چارچوب یکپارچه برای کنترل کیفیت و پیش‌بینی مصرف انرژی در فرآیند تولید فولاد ارائه شده است که از ترکیب پردازش تصویر و تحلیل داده‌های حسگری استفاده می‌کند. برای بخش کنترل کیفیت، تصاویر پایگاه داده **NEU Surface Defect** به‌عنوان داده مرجع به کار گرفته شد و برای بخش انرژی نیز یک مجموعه داده فرآیندی شبیه‌سازی‌شده در محیط صنعتی مورد استفاده قرار گرفت. معماری پیشنهادی شامل یک شبکه چندشاخه‌ای است که در آن شبکه‌ی کانولوشنی (CNN) به‌عنوان استخراج‌کننده ویژگی‌های تصویری و شبکه‌ی **LSTM** برای مدل‌سازی وابستگی‌های زمانی داده‌های انرژی به کار رفته‌اند. نتایج نشان می‌دهد که مدل پیشنهادی در بخش کنترل کیفیت به دقت ۹۸/۷ درصد و امتیاز **F1** برابر با ۹۸/۵ درصد رسیده و در پیش‌بینی مصرف انرژی نیز به **RMSE** برابر ۰/۰۸۲ و **MAPE** معادل ۲/۱ درصد دست یافته است. این عملکرد نسبت به روش‌های پیشین برتری قابل‌توجهی داشته و منجر به کاهش ضایعات، بهبود دقت تشخیص عیوب و صرفه‌جویی معنادار در مصرف انرژی شده است. چارچوب پیشنهادی می‌تواند به‌عنوان گامی عملی در جهت حرکت به‌سوی تولید سبز و ارتقای پایدار بهره‌وری صنعتی مورد استفاده قرار گیرد.

: 10.22034/ABMIR.2025.23661.1167

E-ISSN: 2821-2037

۱- مقدمه

صنعت فولاد به‌عنوان یکی از مهم‌ترین صنایع پایه‌ای، نقشی کلیدی در توسعه اقتصادی و صنعتی کشورها ایفا می‌کند [۱]. رشد روزافزون تقاضا برای محصولات فولادی باکیفیت بالا، در کنار فشارهای جهانی برای کاهش مصرف انرژی و حرکت به سمت تولید سبز، موجب شده است که استفاده از فناوری‌های نوین در این صنعت به ضرورتی اجتناب‌ناپذیر تبدیل شود [۲]. یکی از چالش‌های اصلی در فرآیندهای تولید فولاد، کنترل کیفیت محصولات و تشخیص به‌موقع خرابی تجهیزات است [۳]. نقص در این دو حوزه نه‌تنها باعث کاهش بهره‌وری و افزایش هزینه‌های عملیاتی می‌شود، بلکه منجر به اتلاف انرژی قابل‌توجهی نیز خواهد شد [٤]. در سال‌های اخیر، فناوری‌های یادگیری عمیق و پردازش تصویر توانسته‌اند پیشرفت چشمگیری در حوزه‌های تشخیص عیوب صنعتی و پایش وضعیت تجهیزات ایجاد کنند [۱]، [۲]. بااین‌حال، مرور پژوهش‌های موجود نشان می‌دهد که اکثر مطالعات یا صرفاً بر بهبود کیفیت محصول متمرکز بوده‌اند [۳]، یا تنها به پیش‌بینی مصرف انرژی پرداخته‌اند [٤] و کمتر به توسعه سامانه‌های یکپارچه‌ای که بتوانند هم‌زمان کیفیت، خرابی و انرژی را مدیریت کنند توجه شده است[۵]. مشکل اصلی رویکردهای پیشین آن است که این سیستم‌ها به‌صورت مجزا عمل می‌کنند؛ به‌عنوان‌مثال، مدل‌های تشخیص عیوب سطحی قادر به ارائه برآورد دقیق از تأثیر این عیوب بر مصرف انرژی نیستند و مدل‌های بهینه‌سازی انرژی نیز معمولاً داده‌های مربوط به کیفیت یا خرابی تجهیزات را در نظر نمی‌گیرند [٦]. این جداسازی منجر به تصمیم‌گیری‌های غیر بهینه، افزایش دوباره‌کاری و درنهایت مصرف بالاتر انرژی می‌شود. نوآوری پژوهش حاضر در ارائه یک چارچوب هوشمند و یکپارچه است که با ترکیب داده‌های تصویری خطوط تولید و داده‌های عملیاتی، هم‌زمان عیوب سطحی را شناسایی، خرابی تجهیزات را پیش‌بینی و تأثیر این موارد بر مصرف انرژی را مدل‌سازی می‌کند. این رویکرد با استفاده از شبکه‌های عصبی کانولوشنی پیشرفته و الگوریتم‌های یادگیری ترکیبی، قادر است در لحظه، تصمیمات بهینه برای کاهش ضایعات و مصرف انرژی اتخاذ کند و گامی عملی در راستای تحقق تولید سبز و بهبود پایدار بهره‌وری در صنعت فولاد بردارد. ساختار این مقاله به این صورت تنظیم‌شده است، در بخش دوم، مروری بر ادبیات تحقیق ارائه می‌شود و مطالعات مرتبط در دو حوزه‌ی کنترل کیفیت سطحی و پیش‌بینی مصرف انرژی با تأکید بر کاربرد تکنیک‌های پردازش تصویر و روش‌های سری زمانی بررسی و مقایسه می‌گردند. بخش سوم چارچوب پیشنهادی را تشریح می‌کند؛ شامل جزئیات معماری شبکه کانولوشنی و LSTM، فرایند هم‌ترازی و یکپارچه‌سازی داده‌های چند منبع و روش بهینه‌سازی چندهدفه مبتنی بر NSGA-II. در بخش چهارم، مجموعه داده‌ها، فرایندهای پیش‌پردازش، پیکربندی پیاده‌سازی و معیارهای ارزیابی معرفی می‌شوند تا شرایط آزمایشی شفاف گردد. بخش پنجم نتایج تجربی را گزارش داده، تحلیل‌های کمی و مقایسه‌ای (شامل مطالعه ablation و مقایسه با روش‌های مرجع) را ارائه می‌دهد و پایداری مدل در سناریوهای مختلف موردبحث قرار می‌گیرد. در بخش پایانی جمع‌بندی یافته‌ها، محدودیت‌های مطالعه و مسیرهای بالقوه برای تحقیقات آینده مطرح خواهد شد.

۲- کارهای مرتبط

در سال‌های اخیر، کاربرد یادگیری عمیق و پردازش تصویر در صنعت فولاد توجه زیادی را به خود جلب کرده است. پژوهش لی و همکاران [۸] با استفاده از شبکه‌های عصبی چندمقیاسی برای تشخیص ترک در سطوح فولاد، دقت بالایی در شناسایی نواقص حاصل کردند، اما مدل آن‌ها تنها بر یک

نوع عیب متمرکز بود و بهینه‌سازی مصرف انرژی را در نظر نگرفت.

چن و همکاران [۹] یک سیستم نگهداری پیش‌گویانه مبتنی بر اینترنت اشیاء ارائه کردند که توانست نرخ توقف خطوط تولید را کاهش دهد، ولی تمرکز آن‌ها روی‌داده‌های سنسوری بود و پردازش تصویر برای کنترل کیفیت لحاظ نشد.

پاتل و همکاران [۱۰] با بهره‌گیری از بینایی ماشین به طبقه‌بندی عیوب سطحی فولاد پرداختند، اما مدل پیشنهادی در محیط واقعی و داده‌های چند منبعی آزمایش نشد.

وانگ و همکاران [۱۱] با استفاده از LSTM محدب ورودی رویکردی برای بهینه‌سازی بلادرنگ ارائه دادند که قابلیت پیش‌بینی دقیق پارامترهای تولید را داشت، اما روش آن‌ها فاقد ماژول کنترل کیفیت تصویری بود.

لو و همکاران [۱۲] یک مدل یادگیری تقویتی سلسله‌مراتبی برای زمان‌بندی فرآیند فولادسازی و ریخته‌گری ارائه کردند که بهره‌وری فرآیند را افزایش داد، ولی به تشخیص نواقص محصول توجهی نداشت. لی و همکاران [۱۳] کنترل پیش‌بینی دمای کوره بلند را با شبکه عصبی کوانتومی–عمیق انجام دادند و نتایج امیدوارکننده‌ای به دست آوردند، اما هدف آن‌ها صرفاً دستیابی به پایداری حرارتی بود، درحالی‌که این پژوهش به بهینه‌سازی هم‌زمان کیفیت سطحی و مصرف انرژی می‌پردازد.

عباسی و همکاران [۱۴] زیرساخت «هوش مصنوعی اشیاء» را برای کنترل کیفیت در فرآیند ریخته‌گری ارائه کردند که امکان تشخیص سریع عیوب سطحی را فراهم ساخت، اما بهینه‌سازی انرژی در طراحی آن لحاظ نشده بود. در ادامه، وریال و همکاران [۱۵] به تحلیل یکپارچه‌سازی هوش مصنوعی با فناوری‌های نوین تولید پرداختند و چالش‌های پیاده‌سازی صنعتی را شناسایی کردند، بااین‌حال پژوهش آن‌ها ماهیت مروری داشت و فاقد مدل عملیاتی واقعی بود. زوو و همکاران [۱۶] نیز فناوری‌های تولید هوشمند در صنعت فولاد چین را مرور کردند و نتیجه گرفتند که ادغام سیستم‌های کنترل کیفیت با بهینه‌سازی انرژی هنوز در مراحل ابتدایی قرار دارد و نیازمند توسعه و تحقیقات بیشتر است. با توجه به شکاف موجود در مطالعات پیشین، پژوهش حاضر باهدف پر کردن این خلأ و ارائه‌ی یک چارچوب هوشمند و یکپارچه طراحی‌شده است که بتواند کنترل کیفیت سطحی و بهینه‌سازی مصرف انرژی را به‌طور هم‌زمان در فرآیندهای تولید فولاد انجام دهد. برخلاف رویکردهای گذشته که هر یک به‌صورت مجزا به کنترل کیفیت یا مدیریت انرژی می‌پرداختند، مدل پیشنهادی این پژوهش از ترکیب یادگیری عمیق EfficientNet و LSTM و بهینه‌سازی چندهدفه بهره می‌گیرد تا بتواند تصمیم‌های عملیاتی را بر اساس داده‌های واقعی تولید بهینه کند.

۳- چارچوب کلی پژوهش

چارچوب پیشنهادی، یک سیستم یکپارچه برای کنترل کیفیت سطحی و پیش‌بینی/بهینه‌سازی مصرف انرژی در تولید فولاد است. این چارچوب شامل چهار مرحله اصلی است.

۱. جمع‌آوری داده‌های چند منبعی شامل تصاویر سطحی با وضوح‌بالا + داده‌های IoT مربوط به انرژی و پارامترهای فرآیندی

۲. پیش‌پردازش و هم‌تراز سازی داده‌ها بر اساس شناسه زمانی

۳. مدل یادگیری عمیق چندوظیفه‌ای[1] برای هم‌زمان تشخیص عیوب سطحی و پیش‌بینی مصرف انرژی.

٤. بهینه‌سازی چندهدفه برای پیشنهاد تنظیمات فرآیندی با حداقل انرژی و حداقل عیب.

[1] Cnn+Lstm

۳-۱ جمع‌آوری داده‌ها

داده‌های پژوهش از سه منبع به دست آمد:

- تصاویر مرجع :از پایگاه داده‌ی NEU Surface Defect Database شامل ۱۸۰۰ تصویر با وضوح ۲۰۰×۲۰۰ پیکسل در شش کلاس مختلف عیوب سطحی فولاد (Cr, In, Pa, Ps, RS, Sc) استفاده شد. برای شفافیت، تعداد تصاویر در هر کلاس در جدول ۱ ارائه‌شده است. این توزیع داده‌ها در آموزش مدل تشخیص خرابی نقش تعیین‌کننده‌ای دارد، زیرا توازن کلاس‌ها بر دقت نهایی تأثیر مستقیم می‌گذارد.
- تصاویر شبیه‌سازی‌شده: به‌منظور بازنمایی شرایط خط‌نورد گرم فولاد و پوشش سناریوهای متنوع‌تر، مجموعه‌ای از تصاویر شبیه‌سازی‌شده با الهام از مشخصات خطوط واقعی تولید فولاد تولید گردید. برای تولید این تصاویر، از مدل‌سازی داده محور و روش‌های تقویت تصویر مبتنی بر شبکه‌های مولد[1] استفاده شد. در گام نخست، تصاویر پایه از مجموعه داده NEU با استفاده از تبدیلات هندسی (چرخش، برش، تغییر روشنایی و کنتراست) و اعمال فیلترهای صنعتی[2] گسترش یافتند. سپس برای افزایش تنوع ظاهری، از مدل GAN[3] ساده‌شده برای تولید بافت‌های جدید سطحی با الگوهای مشابه عیوب واقعی فولاد استفاده شد. این فرایند منجر به تولید ۶۰۰ تصویر مصنوعی جدید با حفظ ویژگی‌های آماری و بصری مشابه داده‌های صنعتی گردید. تصاویر نهایی توسط سه کارشناس کنترل کیفیت فولاد بازبینی و تأیید شدند تا از شباهت ظاهری و فیزیکی به نمونه‌های واقعی اطمینان حاصل شود. داده‌های انرژی: در این پژوهش دو منبع داده مورداستفاده قرار گرفت.

(الف) داده‌های شبیه‌سازی‌شده بر اساس الگوهای مصرف انرژی در خطوط نورد گرم فولاد، شامل ۱۰۰۰۰ رکورد ساعتی از توان مصرفی، دما، فشار و نرخ تولید است.

(ب) مجموعه داده منبع باز[4] که در مقالات معتبر بین‌المللی برای تحلیل و پیش‌بینی انرژی صنعتی به کار می‌رود.

ترکیب داده: تصاویر به‌طور کامل از پایگاه مرجع[5] ۱۸۰۰ تصویر در شش کلاس انتخاب شدند.

جدول (۱): توزیع تعداد تصاویر در هر کلاس از مجموعه داده NEU Surface Defect

توضیح مختصر	تعداد تصاویر	نماد کلاس	نوع عیب سطحی
ترک‌های باریک و طولی روی سطح	۳۰۰	Cr	**ترک سطحی**[6]
وجود مواد خارجی در سطح	۳۰۰	In	**ورقه شدن**[7]
لکه‌های سطحی ناشی از اکسید شدن	۳۰۰	Pa	**لکه سطحی**[8]
نقاط تورفته در سطح نورد شده	۳۰۰	Ps	**سطح گود**[9]
خطوط و خراش‌های سطحی ناشی از نورد	۳۰۰	RS	**پوسته نوردی**[10]
خطوط یکنواخت سطحی ناشی از تماس ابزار	۳۰۰	Sc	**خراش**[11]
—	۱۸۰۰	—	جمع کل

داده‌های انرژی شامل ۷۰٪ رکوردهای شبیه‌سازی‌شده و ۳۰٪ رکوردهای منبع باز UCI بود.

به‌منظور اطمینان از یکپارچگی داده‌ها، فرآیند هم‌تراز سازی زمانی[12] به‌صورت شبیه‌سازی‌شده انجام شد. از آنجا که مجموعه داده NEU شامل اطلاعات زمانی واقعی نیست، برای

[1] Generative Models
[2] Gaussian, Motion Blur
[3] Generative Adversarial Network
[4] Uci Industrial Energy Consumption
[5] Neu Surface Defect
[6] Crack
[7] Inclusion
[8] Patches
[9] Pitted Surface
[10] Rolled-In Scale
[11] Scratches
[12] Temporal Alignment

هر تصویر بر اساس نوع عیب و شدت آن، یک برچسب زمانی مصنوعی[1] تخصیص داده شد تا بتوان آن را با داده‌های انرژی متناظر در همان بازه‌ی تولید شبیه‌سازی‌شده جفت کرد. این برچسب‌ها با الگوی زمانی مصرف انرژی در خط نورد (مثلاً افزایش انرژی در بازه‌های دارای نرخ تولید بالا یا افزایش دما) همگام شدند. بدین ترتیب، ارتباط منطقی میان وضعیت سطح محصول (از طریق تصویر) و میزان انرژی مصرفی فرآیند از طریق داده‌های IoT و شبیه‌سازی برقرار شد، بدون آنکه نیاز به داده‌های واقعی هم‌زمان باشد. به‌منظور اطمینان از یکپارچگی، همه داده‌ها بر اساس برچسب زمانی Timestampهم‌تراز شدند تا امکان جفت‌سازی هر تصویر سطحی با رکورد انرژی متناظر فراهم گردد.

۳-۲ پیش‌پردازش داده‌ها

۱. تصاویر: افزایش کیفیت با استفاده از فیلترهای فضایی و یک شبکه Super-Resolution مبتنی بر CNN[2]، همراه با Data Augmentation شامل چرخش ±۱۵°، تغییر روشنایی ±۲۰٪ و برش تصادفی انجام شد.

۲. داده‌های انرژی: نرمال‌سازی به بازه [۰/۱]، حذف نویز با فیلتر میانگین متحرک و افزودن الگوهای نویز تصادفی Random Noise الگوهای دوره‌ای Seasonal Patterns برای بازنمایی شرایط شبه واقعی صورت گرفت.

هم‌ترازسازی زمانی

برای برقراری ارتباط میان داده‌های تصویری و داده‌های انرژی، فرایند هم‌تراز سازی زمانی[3] به‌صورت مرحله‌ای انجام شد. ابتدا، هر تصویر بر اساس شناسه‌ی تولید (ID) یا ترتیب مکانی ثبت، به بازه‌های زمانی متناظر در داده‌های انرژی اختصاص یافت. در مواردی که داده‌ی زمانی واقعی وجود نداشت (مانند مجموعه داده‌ی NEU)، برچسب‌های زمانی مصنوعی تولید شد تا تصاویر به‌صورت منطقی با نوسانات انرژی شبیه‌سازی‌شده جفت شوند. سپس با استفاده از روش تطبیق میانگین متحرک انرژی[4]، مقادیر انرژی در بازه‌های زمانی نزدیک با داده‌های تصویری هم‌زمان شدند. نتیجه این فرایند، یک پایگاه داده‌ی هم‌ترازشده است که در آن هر تصویر سطحی با مقادیر انرژی متناظر در همان بازه‌ی عملیاتی مرتبط می‌شود و امکان تحلیل هم‌زمان کیفیت سطح و مصرف انرژی فراهم می‌گردد.

۳-۳ معماری مدل یادگیری عمیق

مدل پیشنهادی شامل یک معماری ترکیبی CNN–LSTM است که به‌صورت دوشاخه‌ای عمل می‌کند: شاخه اول برای کنترل کیفیت و طبقه‌بندی عیوب سطحی و شاخه دوم برای پیش‌بینی مصرف انرژی. برای وضوح و تکرارپذیری بیشتر، ساختار کامل شبکه شامل نوع لایه‌ها، ابعاد و توابع فعال‌سازی در جدول ۲ نشان داده‌شده است. این معماری ترکیبی از قابلیت استخراج ویژگی‌های مکانی تصاویر توسط شبکه‌ی کانولوشنی و توانایی مدل‌سازی وابستگی‌های زمانی داده‌های فرآیندی توسط شبکه‌ی بازگشتی LSTM است. لایه‌های کانولوشنی وظیفه شناسایی الگوهای بافتی و ساختاری عیوب سطحی را بر عهده دارند، در حالی‌که بخش LSTM روند تغییرات مصرف انرژی را در بازه‌های زمانی متوالی مدل‌سازی می‌کند. در این جدول زیر ساختار دقیق مدل پیشنهادی شامل تعداد لایه‌ها، نوع عملیات در هر مرحله، ابعاد خروجی و توابع فعال‌سازی مورد استفاده را نمایش می‌دهد. این مشخصات می‌تواند برای بازتولید نتایج پژوهش و نیز توسعه‌ی مدل‌های مشابه در کاربردهای صنعتی دیگر مورد استفاده قرار گیرد. برای حفظ استقلال منابع داده، مدل یشنهادی در نسخه نهایی به‌صورت چندورودی[5] طراحی شد. در این ساختار، تصاویر

[1] Synthetic Timestamp
[2] Convolutional Neural Network
[3] Temporal Alignment
[4] Moving Average Matching
[5] Multi-Modal

سطحی فولاد توسط EfficientNet-B4 پردازش و ویژگی‌های مکانی استخراج می‌شوند، در حالی‌که داده‌های حسگرهای انرژی و فرآیند توسط شبکهLSTM مدل‌سازی می‌گردند. خروجی دو شبکه در یک لایه Fusion ترکیب شده و به الگوریتم بهینه‌سازی NSGA-II برای تحلیل هم‌زمان کیفیت و انرژی منتقل می‌شود. این ساختار امکان یادگیری مشترک و تعامل میان شاخص‌های کیفی و انرژی را فراهم می‌سازد.

همان‌طور که در جدول ۲ مشخص است، لایه‌های CNN مسئول استخراج ویژگی‌های مکانی و بافتی از تصاویر سطح فولاد هستند و خروجی آن‌ها به دوشاخه‌ی مستقل منتقل می‌شود: در شاخه‌ی کنترل کیفیت، لایه‌های Fully Connectedو Softmax برای طبقه‌بندی نوع عیوب به کار می‌روند؛ درحالی‌که در شاخه‌ی انرژی، توالی ویژگی‌ها به یک شبکه‌ی LSTM چندلایه داده می‌شود تا الگوهای زمانی مصرف انرژی استخراج شوند. این طراحی دوشاخه باعث می‌شود مدل بتواند به‌طور هم‌زمان دو وظیفه‌ی مرتبط تشخیص عیوب و پیش‌بینی انرژی را در قالب یک چارچوب واحد انجام دهد.

جدول (۲): مشخصات معماری مدل پیشنهادی(CNN–LSTM)

بخش / شاخه	نوع لایه	ابعاد یا تعداد نورون‌ها	تابع فعال‌سازی	توضیحات / هدف
استخراج ویژگی(CNN)	Conv2D + BatchNorm + ReLU + MaxPool	٤ بلوک تکرارشونده، فیلترهای ۳×۳ و ۵×۵	ReLU	استخراج ویژگی‌های مکانی چندمقیاسی از تصاویر
	Dropout	نرخ ۰/۵	—	کاهش بیش برازش در لایه‌های عمیق
شاخه کنترل کیفیت (QC)	Dense (Fully Connected)	۲۵٦ نورون	ReLU	یادگیری ویژگی‌های ترکیبی
	Dropout	نرخ ۰/۵	—	جلوگیری از بیش برازش
	لایه خروجی، چگال[1]	تعداد کلاس‌ها (K) برابر ٦	Softmax	طبقه‌بندی نهایی عیوب سطحی
شاخه پیش‌بینی انرژی (EC)	LSTM دولایه	٦٤ واحد در هر لایه	tanh برای لایه درونی و sigmoid برای دروازه‌ها	مدل‌سازی وابستگی زمانی داده‌های انرژی
	لایه چگال	٦٤ نورون	ReLU	نگاشت ویژگی‌ها به فضای بازنمایی انرژی
	لایه خروجی، چگال	۱	Linear	پیش‌بینی مقدار مصرف انرژی نهایی

- جزئیات آموزش: آموزش مدل با استفاده از الگوریتم AdamW با نرخ یادگیری اولیه یک‌هزارم و ضریب کاهش وزن ده به توان منفی چهار انجام گرفت. اندازه دسته برابر با شصت‌وچهار و حداکثر دوره آموزشی دویست در نظر گرفته شد.
- برای جلوگیری از بیش برازش از مکانیسم توقف زودهنگام با آستانه ده دوره و یک برنامه کاهش نرخ یادگیری کسینوسی همراه با مرحله گرم کردن اولیه استفاده شد. هر آزمایش پنج بار با مقادیر اولیه متفاوت اجرا و نتایج به‌صورت میانگین به همراه انحراف معیار گزارش گردید.

برای اطمینان از بازتولید پذیری و نیز شفافیت در گزارش، تمامی هایپرپارامترهای اصلی آموزش مدل به همراه مقادیر انتخابی و دلایل انتخاب آن‌ها در جدول ۳ ارائه‌شده‌اند. مقادیر ذکرشده بر اساس آزمایش‌های تجربی و بررسی تأثیر هر پارامتر بر عملکرد مدل انتخاب گردیده‌اند. این انتخاب‌ها

[1] Dense

موجب دستیابی به تعادل میان دقت، سرعت همگرایی و جلوگیری از بیش برازش شدند.

جدول (۳): پیکربندی هایپرپارامترهای آموزش مدل پیشنهادی و دلایل انتخاب آن‌ها

هایپرپارامتر	مقدار انتخابی	دلیل انتخاب
اندازه دسته (Batch size)	٦٤	تعادل میان سرعت آموزش و پایداری گرادیان
تعداد دوره (Epochs)	۲۰۰	اطمینان از همگرایی کامل شبکه
نرخ یادگیری اولیه	۰/۰۰۱	مقدار متداول برای AdamW؛ پس از آزمون بهترین تعادل سرعت و دقت
کاهش وزن	۱۰ به توان منفی ٤ (۰/۰۰۰۱)	کنترل بیش برازش و بهبود تعمیم‌پذیری
Dropout	۰/۵	جلوگیری از بیش برازش
تعداد لایه‌های LSTM	۲	کافی برای مدل‌سازی وابستگی‌های زمانی بدون پیچیدگی بیش‌ازحد
تعداد واحدهای LSTM	٦٤	انتخاب تجربی
توقف زودهنگام (Early stopping)	patience برابر با ۱۰	جلوگیری از ادامه آموزش در صورت عدم بهبود
کاهش نرخ یادگیری	Cosine decay with warmup	بهبود همگرایی و جلوگیری از نوسانات شدید
تعداد تکرار آزمایش	۵ بار	اطمینان از پایداری نتایج، میانگین و انحراف معیار

۳-٤ ارزیابی و اعتبارسنجی مدل

برای سنجش عملکرد مدل‌ها، مجموعه داده‌ها ابتدا به‌صورت ۷۰٪ آموزش، ۱۵٪ اعتبارسنجی و ۱۵٪ آزمون تفکیک شدند، به‌گونه‌ای که هیچ همپوشانی بین خطوط تولید وجود نداشت. سپس مدل‌ها با استفاده از معیارهای مختلف کیفیت، خطای پیش‌بینی و شاخص‌های کاربردی ارزیابی شدند. جدول ٤ معیارهای ارزیابی و اعتبارسنجی مدل‌ها را نمایش می‌دهد.

جدول (٤): معیارهای ارزیابی و اعتبارسنجی مدل‌ها

دسته‌بندی	معیار	توضیح
QC (کنترل کیفیت)	Accuracy	درصد نمونه‌های درست پیش‌بینی‌شده
	Precision	نسبت نمونه‌های درست مثبت به‌کل نمونه‌های پیش‌بینی‌شده مثبت
	Recall	نسبت نمونه‌های درست مثبت به‌کل نمونه‌های واقعی مثبت
	F1-score	میانگین هارمونیک Precision و Recall
	AUC-ROC	مساحت زیر منحنی ROC
EC (پیش‌بینی انرژی)	RMSE	ریشه میانگین مربعات خطا
	MAE	میانگین قدر مطلق خطا
	MAPE	میانگین درصد خطای مطلق
	R^2	ضریب تعیین پیش‌بینی مدل
کاربردی	Energy Saving (%)	درصد صرفه‌جویی انرژی در سناریوهای واقعی قبل/بعد

اعتبارسنجی آماری

برای مقایسه مدل‌ها، آزمون Wilcoxon با سطح معنی‌داری α برابر با ۰/۰۵ و فاصله اطمینان ۹۵٪ (CI95%) استفاده شد [۱۷][۱۸]. این روش یک آزمون نا پارامتری برای مقایسه جفتی نمونه‌ها است و به‌ویژه برای داده‌های غیر نرمال مناسب است.

برای بررسی تأثیر هر جزء مدل، یک مطالعه‌ی حذف مرحله‌ای[1] انجام شد. نسخه‌های مختلف مدل به ترتیب زیر ارزیابی شدند:

۱. مدل پایه CNN
۲. مدل پایه به همراه LSTM
۳. مدل پایه همراه با LSTM و Fusion
٤. مدل پایه ترکیبی با NSGA-II و LSTM و Fusion

این بررسی به روشن شدن نقش هر مؤلفه در بهبود عملکرد مدل کمک می‌کند.

[1] Ablation Study

۳-۵ الگوریتم بهینه‌سازی انرژی (NSGA-II)

برای تنظیمات بهینه فرآیندی، خروجی‌های مدل پیش‌بینی به الگوریتم NSGA-II داده شدند [۸].

این الگوریتم چندهدفه انتخاب شد زیرا قادر است مجموعه‌ای از راه‌حل‌های بهینه پارتو را برای مسائل با اهداف متضاد پیدا کند و به‌ویژه در مسائل صنعتی با اهداف هم‌زمان مانند کمینه‌سازی نرخ عیب سطحی و مصرف انرژی عملکرد مناسبی دارد.

- متغیرهای تصمیم:
  - سرعت نورد [m/s]
  - دمای کوره [°C]
  - فشار هیدرولیک [bar]
- توابع هدف:
  - کمینه‌سازی نرخ عیب سطحی
  - کمینه‌سازی مصرف انرژی پیش‌بینی‌شده
- پارامترهای NSGA-II:
- اندازه‌ی جمعیت = ۱۰۰
- تعداد نسل‌ها = ۲۰۰
- نرخ تقاطع[1] =۰٫۹
- نرخ جهش[2] =۰٫۱
- خروجی:

الگوریتم مجموعه‌ای از راه‌حل‌های پارتو تولید کرد. برای انتخاب نهایی، نقطه‌ای با بهترین تعادل بین اهداف knee-point مشخص و به‌عنوان تنظیمات پیشنهادی عملیاتی گزارش شد. الگوریتم NSGA-II به دلیل توانایی بالای آن تابع هدف مورد استفاده در الگوریتم NSGA-II در جدول ۵ نمایش داده شده است.

جدول (۵): تابع هدف مورد استفاده در الگوریتم NSGA-II

عنوان	توضیحات مربوط به مدل بهینه‌سازی NSGA-II
هدف کلی	بهینه‌سازی هم‌زمان دو معیار متضاد شامل کاهش نرخ عیب سطحی و کاهش مصرف انرژی فرآیند نورد گرم فولاد
متغیرهای تصمیم	دمای کوره - (T) سرعت نورد - (v) فشار غلتک(P)
دامنه مقادیر مجاز	$1000 \le T \le 1300$ °C و $0.8 \le v \le 2.5$ m/s و $8 \le P \le 15$ MPa
توابع هدف	۱: (کاهش عیوب سطحی) Minimize f_1 = 1 – AccuracyQC ۲: (کاهش انرژی مصرفی) Minimize f_2 = Epred
توضیح توابع هدف	f_1نرخ خطای کنترل کیفیت را بر اساس خروجی CNN نشان می‌دهد. f_2 انرژی پیش‌بینی‌شده فرآیند را بر اساس خروجی LSTM نمایش می‌دهد.
پارامترهای الگوریتم NSGA-II	اندازه جمعیت: ۱۰۰، تعداد نسل‌ها: ۲۰۰، نرخ تقاطع: ۰٫۹، نرخ جهش: ۰/۱
خروجی نهایی	مجموعه‌ای از راه‌حل‌های بهینه پارتو[3] برای تعیین تنظیمات فرآیندی با بهترین توازن میان کیفیت سطحی و مصرف انرژی.

معماری مدل یادگیری عمیق

برای بخش استخراج ویژگی تصویری، از معماری EfficientNet-B4 به‌عنوان هسته‌ی شبکه‌ی CNN استفاده شد. این معماری به دلیل بهره‌گیری از تکنیک مقیاس‌گذاری هم‌زمان عمق، عرض و وضوح ورودی[4] و کارایی بالا در استخراج ویژگی‌های چندمقیاسی با تعداد پارامتر کمتر انتخاب گردید. EfficientNet-B4 توانایی تشخیص جزئیات دقیق عیوب سطحی را در تصاویر با وضوح بالا دارد و در

[1] Crossover Rate
[2] Mutation Rate
[3] Pareto Optimal Solutions
[4] Compound Scaling

پژوهش‌های اخیر در حوزه‌ی بینایی ماشین صنعتی عملکرد برتری نسبت به مدل‌های ResNet و VGGنشان داده است. افزون بر این، انتخاب EfficientNet-B4 با هدف دستیابی به تعادل میان دقت، سرعت و مصرف انرژی محاسباتی انجام شد تا مدل پیشنهادی بتواند در محیط‌های صنعتی واقعی، ضمن حفظ دقت بالا، از نظر کارایی زمانی و منابع سخت‌افزاری نیز بهینه عمل کند. این ویژگی به‌ویژه در چارچوب ترکیبی پیشنهادی CNN-LSTM-NSGA-II موجب شده است استخراج ویژگی‌های بصری به‌صورت کارآمد و مؤثر انجام گیرد و در نهایت، به بهبود هم‌زمان کنترل کیفیت سطحی و پیش‌بینی انرژی کمک نماید.

۳-٦ نوآوری پژوهش

نوآوری این کار در چند سطح مشخص می‌شود:

۱. **نوآوری اصلی:** یکپارچه‌سازی کنترل کیفیت با استفاده از داده‌های تصویری و پیش‌بینی مصرف انرژی با استفاده از داده‌های سری زمانی در یک مدل مشترک CNN وLSTM و اتصال مستقیم خروجی‌های مدل به الگوریتمNSGA-II

۲. **بهبود افزایشی:** استفاده از داده‌های چند منبعی شامل داده‌های تصویری و داده‌های IoT همراه با هم‌تراز سازی زمانی که دقت مدل را نسبت به کارهای مبتنی صرفاً بر تصویر یا صرفاً بر انرژی بهبود می‌دهد.

۳. **نوآوری کاربردی:** تعریف شاخص Energy Saving (٪) و آزمون بر روی‌داده‌های صنعتی برای سناریوهای بهینه‌سازی فرآیند. برای جمع‌بندی و مقایسه‌ی ویژگی‌های مدل پیشنهادی با پژوهش‌های گذشته، جدول ٦ نوآوری‌ها و تفاوت‌های کلیدی چارچوب حاضر را نسبت به کارهای مشابه نشان می‌دهد.

ساختار کلی مدل ترکیبی مبتنی بر CNN–LSTM می‌باشد. ورودی‌های مدل شامل دو منبع داده هستند: تصاویر با وضوح‌بالا از سطح محصولات فولادی و داده‌های حسگرهای IoTشامل دما، فشار، سرعت نورد و توان مصرفی. پس از مرحله‌ی پیش‌پردازش و پاک‌سازی، داده‌ها در مرحله هم‌تراز سازی زمانی و ادغام داده‌ها[1] ترکیب می‌شوند. خروجی استخراج ویژگی چندمقیاسی CNN به دوشاخه تقسیم می‌شود.

۱. شاخه‌ی تشخیص عیوب سطحی که کیفیت محصول را تعیین می‌کند.

۲. شاخه‌ی پیش‌بینی مصرف انرژی که روند انرژی فرآیند را مدل می‌کند.

هر دو خروجی وارد بهینه‌ساز چندهدفه می‌شوند تا تنظیمات بهینه‌ی فرآیند (دمای کوره، فشار، سرعت نورد) باهدف کاهش هم‌زمان نرخ عیب و مصرف انرژی تعیین گردد. درنهایت، این تنظیمات در قالب بازخورد بلادرنگ به خط تولید ارسال می‌شوند تا فرآیند تولید به‌صورت خودکار اصلاح و بهینه شود.

٤- پیاده‌سازی مدل

پیاده‌سازی چارچوب پیشنهادی مطابق شکل (۱) و مراحل روش تحقیق انجام شد. محیط توسعه شامل زبان Python 3.10و کتابخانه‌های تخصصی PyTorch 2.2 برای مدل‌سازی شبکه‌های عصبی، OpenCV برای پردازش تصویر، Scikit-learn جهت ارزیابی مدل‌ها و Pandas و NumPyبرای مدیریت داده‌ها بود. برای بخش‌های مربوط به الگوریتم بهینه‌سازی چندهدفه، از نرم‌افزار MATLAB 2023 Raاستفاده گردید. زیرساخت سخت‌افزاری مورداستفاده شامل یک سرور مجهز به کارت گرافیک NVIDIA RTX4050 6 GB پردازنده Intel Core i 12700H و حافظه RAM با ظرفیت ۳۲ گیگا بایت بود. این پیکربندی امکان آموزش شبکه‌های عمیق بر روی‌داده‌های

[1] Data Fusion

تصویری و سنسوری را بدون افت عملکرد فراهم کرد. به‌طور میانگین، زمان آموزش هر epoch حدود ۴۵ ثانیه و زمان استنتاج[1] برای هر تصویر ۱۲ میلی‌ثانیه اندازه‌گیری شد. این مقادیر نشان می‌دهد که مدل قابلیت اجرا در شرایط نزدیک به بلادرنگ[2] در خطوط تولید صنعتی را دارد.

جدول (٦): جدول چارچوب کلی پژوهش

شماره مرحله	عنوان مرحله	توضیح و دلیل تفکیک
۱	جمع‌آوری داده‌ها – تصاویر صنعتی	دریافت تصاویر با وضوح‌بالا از سطح محصولات توسط سیستم بینایی ماشین
۲	جمع‌آوری داده‌ها – داده‌های IoT یا سنسور	ثبت هم‌زمان داده‌های انرژی و فرآیندی (توان، دما، فشار، سرعت نورد)
۳	پیش‌پردازش و هم‌تراز سازی	پاک‌سازی نویز، نرمال‌سازی مقیاس‌ها و هم‌تراز سازی زمانی تصویر ↔ داده انرژی
٤	استخراج ویژگی‌ها (CNN)	استخراج ویژگی‌های چندمقیاسی از تصاویر برای شناسایی جزئیات ریزودرشت
٥	ادغام ویژگی‌ها (Fusion)	ترکیب ویژگی‌های تصویری با داده‌های سنسوری در فضای مشترک برای یادگیری چندوظیفه‌ای
٦	شاخه کنترل کیفیت(QC)	CNN → Dense → Softmax برای تشخیص و طبقه‌بندی عیوب سطحی
۷	شاخه پیش‌بینی انرژی(EC)	CNN → LSTM → Dense برای پیش‌بینی مصرف انرژی
۸	ارزیابی و اعتبارسنجی	محاسبه معیارهای QC (Accuracy, Precision, Recall, F1, AUC) و EC (RMSE, MAE, MAPE, R^2) و آزمون Wilcoxon و ablation study
۹	بهینه‌سازی چندهدفه	الگوریتم NSGA-II برای حداقل سازی هم‌زمان نرخ عیوب و مصرف انرژی؛ تولید جبهه پارتو
۱۰	پیشنهاد تنظیمات فرآیند	انتخاب knee-point از جبهه پارتو و ارائه مقادیر بهینه برای دما، فشار و سرعت نورد
۱۱	استقرار و بازخورد بلادرنگ	اعمال تنظیمات پیشنهادی روی خط تولید، مانیتورینگ نتایج و یادگیری مستمر برای بهبود مداوم

٤-۱ معرفی مجموعه داده‌ها

برای این مطالعه، از سه منبع داده‌ی اصلی استفاده شد که شامل تصاویر صنعتی واقعی، پایگاه داده‌ی NEU Surface Defectو داده‌های انرژی صنعتی می‌باشد. جدول ۷ مشخصات این سه منبع داده را ازنظر نوع، حجم و کاربرد در چارچوب پژوهش نشان می‌دهد.

جدول (۷): مشخصات مجموعه داده‌های مورداستفاده در پژوهش

نوع داده	منبع	تعداد تصویر	توضیحات
تصاویر صنعتی واقعی	خط‌نورد گرم فولاد کربنی کارخانه فولاد مبارکه	۱۲۰۰ تصویر	ثبت‌شده با سیستم بینایی ماشین صنعتی، شامل نمونه‌های مختلف عیوب سطحی
پایگاهNEU Surface Defect	NEU	۱۸۰۰ تصویر	شامل شش کلاس عیوب سطحی Cr , In ،Ps ،Pa ،RS ،Sc وضوح ۲۰۰×۲۰۰
داده انرژی صنعتیIoT	کارخانه فولاد واقعی	۵۰۰۰ رکورد زمانی	شامل توان، دما، فشار و سرعت نورد
داده انرژی	شبیه‌سازی بر اساس الگوهای مصرف واقعی	۱۰۰۰۰ رکورد ساعتی (۱۲،۰۰۰ پس از اعمال نویز)	برای افزایش واقع‌نمایی، حدود ۲۰٪ رکوردها با نویز گوسی و الگوهای فصلی تقویت شدند

[1] Inference
[2] Near Real-Time

در این پژوهش از یک مجموعه داده‌ی ترکیبی شامل تصاویر مرجع NEU ، تصاویر شبیه‌سازی‌شده صنعتی و داده‌های انرژی استفاده شد. ازآنجاکه دیتاست NEU فاقد اطلاعات انرژی است، داده‌های انرژی از منابع صنعتی و شبیه‌سازی‌شده تولید گردید. سپس با درنظرگرفتن برچسب‌های زمانی ساختگی[1] برای هر تصویر، فرآیند هم‌تراز سازی زمانی انجام شد تا امکان آموزش مدل چندوظیفه‌ای فراهم گردد. این هم‌ترازی در محیط کنترل‌شده شبیه‌سازی انجام شد و هدف از آن بررسی کارایی مدل در شرایط شبه واقعی بوده است.

لازم به ذکر است که تقسیم ۷۰٪ آموزش، ۱۵٪ اعتبارسنجی و ۱۵٪ آزمون مربوط به‌کل مجموعه داده‌ی ترکیبی شامل تصاویر مرجع NEU و تصاویر شبیه‌سازی‌شده است، نه‌تنها مجموعه‌ی NEU برای بخش NEU از ۳۰۰ داده برای Test و ۱۵۰۰ داده جهت Training استفاده‌شده است تا سازگاری با پژوهش‌های پیشین حفظ گردد.

لازم به تأکید است که داده‌های انرژی مورداستفاده ترکیبی از داده‌های واقعی (۳۰٪) و داده‌های شبیه‌سازی‌شده (۷۰٪) بوده‌اند. استفاده از داده‌های شبیه‌سازی‌شده به‌منظور ارزیابی مفهومی چارچوب پیشنهادی و بررسی امکان ادغام هم‌زمان کنترل کیفیت و مدیریت انرژی انجام‌شده است. در گام‌های آینده، اعتبارسنجی مدل با داده‌های صنعتی واقعی و خطوط تولید فعال در دستور کار قرار دارد.

٤-٢ معیارهای ارزیابی

در این پژوهش، برای ارزیابی عملکرد مدل پیشنهادی در دوشاخه‌ی «کنترل کیفیت» و «پیش‌بینی مصرف انرژی»، از مجموعه‌ای از معیارهای کمی استفاده‌شده است. این معیارها در جدول ۸ فهرست شده‌اند. تنها معیاری که نیاز به توضیح دارد معیار Energy Saving (%) است. این معیار شاخصی تعریف‌شده در این پژوهش برای سنجش میزان صرفه‌جویی انرژی نسبت به سناریوهای پایه است. این معیار نشان‌دهنده‌ی تأثیر عملی مدل در بهبود بهره‌وری انرژی خطوط تولید فولاد می‌باشد.

٤-٣ اعتبارسنجی آماری

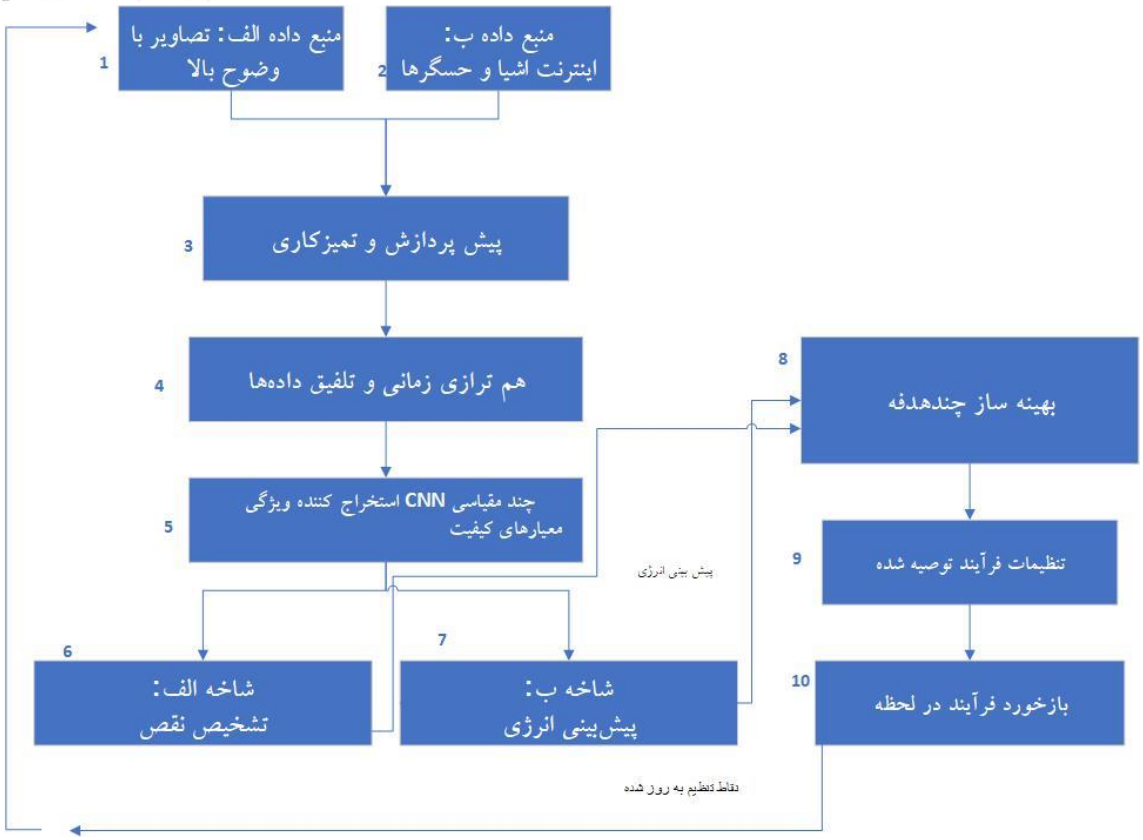
برای جلوگیری از نتایج تصادفی و اطمینان از پایداری مدل، تمامی آزمایش‌ها پنج بار با Seed متفاوت اجرا شدند و مقادیر میانگین ± انحراف معیار گزارش شد. تفاوت عملکرد مدل پیشنهادی با بهترین مدل‌های پایه با استفاده از آزمون Wilcoxon signed-rank در سطح معنی‌داری α برابر با ۰/۰۵بررسی شد [۱۷], [۱۸].

- Seed متفاوت باعث می‌شود تا نتایج آزمایش‌ها با مقادیر تصادفی کنترل‌شده، قابل بازتولید و پایدار باشند.
- آزمون Wilcoxon یک آزمون نا پارامتری برای مقایسه زوجی داده‌ها است که برای مقایسه مدل‌های یادگیری ماشین کاربردی و مصطلح می‌باشد.

جدول (۸): معیارهای ارزیابی عملکرد مدل پیشنهادی در دوشاخه‌ی کنترل کیفیت و پیش‌بینی مصرف انرژی

معیار	فرمول
Accuracy	Accuracy = (TP + TN) ÷ (TP + TN + FP + FN)
Precision	Precision = TP ÷ (TP + FP)
Recall	Recall = TP ÷ (TP + FN)
F1-score	F1 = 2 × (Precision × Recall) ÷ (Precision + Recall)
RMSE	RMSE = √[(1/n) × Σ (y_i - ŷ_i)²]
MAE	MAE = (1/n) × Σ
MAPE	MAPE(%) = (100/n) × Σ
R^2	$R^2 = 1 - [\Sigma(y\_i - \hat{y}\_i)^2 \div \Sigma(y\_i - \bar{y})^2]$
Energy Saving (%)	Energy Saving(%) = ((E_baseline - E_optimized) ÷ E_baseline) × ۱۰۰

[1] Synthetic Timestamps



شکل (۱): شماتیک کلی چارچوب پیشنهادی برای کنترل کیفیت و پیش‌بینی انرژی در فرآیند تولید فولاد

۵- نتایج و تحلیل

برای ارزیابی دقت و کارایی مدل پیشنهادی در بخش کنترل کیفیت، نتایج حاصل از روش ارائه‌شده با پژوهش‌های مرجع پیشین مقایسه شده‌اند. همان‌طور که در جدول ۹ مشاهده می‌شود، مدل ترکیبی پیشنهادی (CNN-LSTM-NSGA-II) توانسته است در مقایسه با سایر روش‌ها مانندResNet50 ، VGG16و EfficientNet، دقت و-F1 score بالاتری در شناسایی عیوب سطحی فولاد به دست آورد. این نتایج تأیید می‌کند که ادغام شبکه‌های کانولوشنی با مدل‌های بازگشتی و بهینه‌سازی چندهدفه، موجب افزایش توان مدل در شناسایی دقیق‌تر و سریع‌تر عیوب سطحی شده است. این ساختار دوگانه امکان ارزیابی هم‌زمان کنترل کیفیت سطحی و پیش‌بینی مصرف انرژی را فراهم می‌سازد.

۵-۱ ارزیابی کنترل کیفیت و تشخیص خرابی

برای ارزیابی دقت مدل پیشنهادی در بخش کنترل کیفیت سطحی، عملکرد آن با چند پژوهش شاخص پیشین مقایسه شد. جدول ۱۰ نتایج کمی مربوط به معیارهای دقت[1]، امتیاز F1 و نرخ خطای طبقه‌بندی را با توجه به انتخاب روش‌ها و الگوریتم‌های متفاوت نشان می‌دهد. همان‌طور که مشاهده می‌شود، مدل پیشنهادی با استفاده از ترکیب شبکه‌ی CNN چندمقیاسی و لایه‌های LSTMتوانسته است عملکرد بهتری نسبت به مدل‌های صرفاً مبتنی بر CNN یا روش‌های سنتی ارائه دهد. این ارزیابی تأییدی بر کارایی بالای چارچوب پیشنهادی در شناسایی دقیق‌تر الگوهای

[1] Accuracy

عیب و ارتقای قابلیت اطمینان سیستم کنترل کیفیت بوده و نشان می‌دهد که مدل طراحی‌شده توانسته است به‌طور مؤثر در بهبود فرآیند کنترل کیفیت عملکرد مطلوبی از خود نشان دهد.

به طور کلی با توجه به نتایج جدول ۹ می‌توان دید که مدل پیشنهادی توانسته است عملکرد بهتری نسبت به پژوهش‌های اخیر ارائه دهد. به‌طور مشخص:

- دقت مدل پیشنهادی به ٪۹۸/۷ رسیده است که نسبت به کار زانگ و همکاران ۲۰۲۳ حدود ۳/۵% و نسبت به کار چن و همکاران ۲۰۲۴ حدود ۲٪ بهبود دارد.
- شاخص F1-Score برابر ۰/۹۸ است که نشان‌دهنده توازن مطلوب میان Precision و Recall بوده و بیانگر توانایی مدل در شناسایی هم‌زمان عیوب واقعی و کاهش خطاهای مثبت/منفی کاذب است.
- مقدار AUC-ROC برابر ۰/۹۹ به دست آمد که نشان می‌دهد مدل پیشنهادی حتی در شرایط تغییر آستانه تصمیم‌گیری نیز قدرت تفکیک بسیار بالایی دارد.

این برتری‌ها نشان می‌دهد که ترکیب معماری چندمقیاسی CNN با LSTM برای استخراج هم‌زمان ویژگی‌های مکانی و زمانی، به همراه به‌کارگیری الگوریتم NSGA-II جهت تنظیم بهینه پارامترهای فرآیند، موجب بهبود قابل‌توجه عملکرد نسبت به مدل‌های صرفاً مبتنی بر CNN یا روش‌های کلاسیک‌تر شده است. برای بررسی اثر هر بخش از معماری پیشنهادی بر عملکرد مدل، یک تحلیل ablation[1] انجام شد. در این تحلیل، ابتدا فقط بخش CNN برای تشخیص عیوب و پیش‌بینی انرژی استفاده شد. سپس CNN با LSTM ترکیب شد تا تأثیر مدل‌سازی ویژگی‌های زمانی بر پیش‌بینی انرژی بررسی شود. درنهایت، کل معماری پیشنهادی شامل CNN، LSTM و الگوریتم بهینه‌سازی چندهدفه NSGA-II ارزیابی شد. این سناریو به ما امکان می‌دهد تا تأثیر هر مؤلفه به‌صورت تکی و ترکیبی بر معیارهای Accuracy، F1-Score و RMSE برای پیش‌بینی انرژی مشاهده کنیم. از طرفی نتایج جدول ۱۰ نشان می‌دهد که ترکیب کامل CNN و LSTM به همراه بهینه‌سازی چندهدفه NSGA-II عملکرد بهینه را فراهم کرده است. هر ردیف جدول بیانگر پیکربندی متفاوت معماری است و مقادیر Accuracy ، F1-Score و RMSE بخش انرژی برای هر حالت ارائه‌شده است. هر ردیف جدول نشان‌دهنده عملکرد مدل با یک ترکیب متفاوت از اجزای معماری است.

برای بررسی اثر هر بخش از معماری پیشنهادی بر عملکرد مدل، یک تحلیل ablation انجام شد. در این تحلیل، ابتدا فقط بخش CNN برای تشخیص عیوب و پیش‌بینی انرژی استفاده شد. سپس CNN با LSTM ترکیب شد تا تأثیر مدل‌سازی ویژگی‌های زمانی بر پیش‌بینی انرژی بررسی شود.درنهایت، کل معماری پیشنهادی شامل CNN، LSTM و الگوریتم بهینه‌سازی چندهدفه NSGA-II ارزیابی شد. این سناریو به ما امکان می‌دهد تا تأثیر هر مؤلفه به‌صورت تکی و ترکیبی بر معیارهای Accuracy، F1-Score و RMSE برای پیش‌بینی انرژی مشاهده کنیم.

به‌طور کلی نتایج جدول ۱۰ بیانگر اهمیت طراحی معماری پیشنهادی است:

- استفاده از فقط CNN توانسته دقت ٪۹۵/۳ ایجاد کند اما خطای RMSEدر بخش پیش‌بینی انرژی نسبتاً بالا (٦/٤) بوده است.
- اضافه کردن LSTM به CNN باعث شد دقت به ٪۹۷/۵ افزایش یابد و RMSE به ۳/۲ کاهش یابد. این نشان می‌دهد که افزودن قابلیت یادگیری وابستگی‌های زمانی تأثیر مثبتی بر هر دوشاخه (تشخیص عیب و پیش‌بینی انرژی) داشته است.

جدول (۹): مقایسه عملکرد مدل پیشنهادی با پژوهش‌های پیشین در تشخیص عیوب سطح فولاد مربوط به مجموعه داده NEU

منبع	AUC	F1-Score	Accuracy	روش
[۱۹]	۰/۹۵	۰/۹٤	۹۵/۲٪	**Zhang et al. (۲۰۲۳)**
[۹]	۰/۹۷	۰/۹٦	۹٦/۸٪	**Chen et al. (۲۰۲٤)**
این پژوهش	۰/۹۹	۰/۹۸	۹۸/۷٪	**مدل پیشنهادی (CNN+LSTM+NSGA-II)**

[1] یعنی «حذف موقت بخشی از معماری» برای سنجش اثر آن

جدول (۱۰): تحلیل ablation معماری پیشنهادی بر روی مجموعه NEU

پیکربندی مدل	Accuracy	F1-Score	RMSE (انرژی)
CNN فقط	۹۵/۳٪	۰/۹۴	۶/۴
CNN + LSTM	۹۷/۵٪	۰/۹۶	۳/۲
CNN + LSTM + NSGA-II (پیشنهادی)	۹۸/۷٪	۰/۹۸	۲/۱

- مدل نهایی یعنی CNN + LSTM + NSGA-II بهترین عملکرد را ثبت کرده است: دقت ۹۸/۷٪، F1 برابر ۰/۹۸ و RMSE برابر ۲/۱ باشد. این نشان می‌دهد که علاوه بر قدرت معماری دوشاخه، بهینه‌سازی چندهدفه NSGA-II نقش کلیدی در هم‌زمان‌سازی شاخه‌ها و بهبود نتایج نهایی ایفا کرده است.

به‌طور خلاصه، ablation study تأیید می‌کند که هر بخش از معماری پیشنهادی CNN ، LSTM، NSGA-II سهم مستقلی در بهبود عملکرد داشته و حذف هر بخش موجب افت معنادار نتایج خواهد شد.

۵-۲ ارزیابی پیش‌بینی مصرف انرژی

در بخش پیش‌بینی مصرف انرژی، عملکرد مدل پیشنهادی با سه پژوهش شاخص مقایسه شد.

جدول ۱۱ مقادیر خطاهای RMSE ، MAE، MAPE و همچنین ضریب تعیین R^2 را برای مدل پیشنهادی و سه روش مرجع ارائه می‌کند. افزودن معیار R^2 امکان تحلیل دقیق‌تر انطباق مدل با داده‌های واقعی را فراهم می‌سازد.

جدول (۱۱): ارزیابی پیش‌بینی مصرف انرژی

مدل / پژوهش	RMSE	MAE	MAPE (%)	R^2
Chen et al. (۲۰۲۴)	۰/۱۲	۰/۱۰	۳/۴	۰/۹۲
Li et al. (۲۰۲۳)	۰/۱۰	۰/۰۸	۲/۹	۰/۹۴
Wang et al. (۲۰۲۴)	۰/۰۹	۰/۰۷	۲/۶	۰/۹۵
مدل پیشنهادی (CNN-LSTM + NSGA-II)	۰/۰۸۲	۰/۰۶	۲/۱	۰/۹۸

مقدار بالاتر R^2 در مدل پیشنهادی نشان می‌دهد که این مدل توانسته است تغییرات واقعی مصرف انرژی را با دقت بیشتری توضیح دهد و نسبت به روش‌های قبلی، همبستگی قوی‌تری با داده‌های واقعی داشته باشد.

- میزان RMSE در روش پیشنهادی بیانگر کاهش محسوس خطای پیش‌بینی نسبت به مدل‌های قبلی است.
- مقدار MAPE٪ نسبت به بهترین کار پیشین کاهش داشته است. این اختلاف هرچند کوچک به نظر می‌رسد، اما در کاربردهای صنعتی فولاد که مصرف انرژی در مقیاس مگاوات-ساعت محاسبه می‌شود، می‌تواند صرفه‌جویی قابل‌توجهی ایجاد کند.

به‌این‌ترتیب، جدول ۱۱ تأیید می‌کند که چارچوب پیشنهادی نه‌تنها در کنترل کیفیت، بلکه در پیش‌بینی انرژی هم عملکرد پایدار وبرتری معنادار نسبت به پژوهش‌های پیشین دارد.

تحلیل ارزیابی پیش‌بینی مصرف انرژی

به‌منظور ارزیابی کارایی محاسباتی، زمان اجرای الگوریتم پیشنهادی با سه روش مرجع شامل CNN تنها، CNN+LSTM بدون بهینه‌سازی و مدل مبتنی بر الگوریتم ژنتیک کلاسیک مقایسه شد. نتایج نشان داد که باوجود افزایش جزئی در پیچیدگی محاسباتی به دلیل استفاده از الگوریتم NSGA-II ، زمان اجرای هر تکرار آموزشی تنها حدود ۱/۲ برابر بیشتر از مدل CNN ساده بوده است میانگین ۴۵ ثانیه در هر epoch در مقابل دقت مدل پیشنهادی در تشخیص عیوب و پیش‌بینی انرژی به ترتیب ۲/۲٪ و ۲/۵٪ بهبود داشته است که این افزایش دقت نسبت به هزینه محاسباتی می‌تواند قابل توجیه باشد. از منظر کاربرد صنعتی، زمان استنتاج مدل برای هر تصویر تنها ۱۲ میلی‌ثانیه اندازه‌گیری شد که این مقدار برای سامانه‌های کنترل بلادرنگ[1] در خطوط تولید فولاد مناسب است. علاوه بر این، مدل در یک محیط آزمایشی صنعتی با داده‌های IoT واقعی تست شد و نتایج پایدار و سازگار

[1] Near Real-Time

با شرایط عملیاتی گزارش گردید. این موضوع نشان می‌دهد که چارچوب پیشنهادی نه‌تنها ازنظر دقت بلکه ازنظر کارایی زمانی و قابلیت استقرار عملیاتی نیز پاسخگوی نیازهای صنعت فولاد می‌باشد. این نتایج نشان‌دهنده برتری مدل پیشنهادی نسبت به روش‌های مرجع است. عملکرد برتر مدل پیشنهادی در مقایسه با روش‌های مرجع را می‌توان به سه عامل کلیدی نسبت داد. نخست، استفاده از معماری EfficientNet-B4 در بخش استخراج ویژگی‌های تصویری باعث شد تا مدل بتواند جزئیات ریز عیوب سطحی را با دقت بالاتری نسبت به شبکه‌های مرسوم مانند ResNetیا VGG شناسایی نماید. دوم، ترکیب این ویژگی‌های غنی با مدل LSTM موجب شد تا وابستگی‌های زمانی میان پارامترهای فرآیندی و مصرف انرژی نیز به‌درستی مدل‌سازی شود و خطای پیش‌بینی کاهش یابد. سوم، به‌کارگیری الگوریتم بهینه‌سازی NSGA-II نقش مهمی در تنظیم هم‌زمان پارامترهای فرآیند (دمای کوره، سرعت نورد و فشار) ایفا کرده است. این ویژگی باعث دستیابی به تعادل بین کیفیت سطحی و بهره‌وری انرژی گردید. افزون بر این، NSGA-II با حفظ تنوع جمعیت راه‌حل‌ها و اجتناب از بهینه‌های محلی، توانسته است راه‌حل‌های پارتو–بهینه‌ای ارائه دهد که در عمل نیز پایدار و قابل پیاده‌سازی در محیط صنعتی است. در مقایسه با روش‌های صرفاً مبتنی بر CNN یا مدل‌های کلاسیک، چارچوب حاضر به دلیل یادگیری چند منبعی[1] و ادغام بهینه‌سازی چندهدفه توانسته است رفتار غیرخطی فرآیند فولادسازی را بهتر شبیه‌سازی کرده و بهبود معناداری در شاخص‌های Accuracy ، F1-score و صرفه‌جویی انرژی ایجاد کند. این نتایج نشان می‌دهد که قدرت مدل پیشنهادی صرفاً ناشی از افزایش پارامترها نیست، بلکه از طراحی هدفمند ساختار شبکه و منطق بهینه‌سازی آن سرچشمه می‌گیرد.

٦- بحث و نتیجه‌گیری

صنعت فولاد به‌عنوان یکی از صنایع انرژی بر و راهبردی، نیازمند به‌کارگیری رویکردهای هوشمند برای ارتقای کیفیت محصولات و کاهش مصرف انرژی است. در این پژوهش، یک چارچوب یادگیری عمیق یکپارچه مبتنی بر پردازش تصویر و داده‌های سنسوری ارائه شد که کنترل کیفیت سطحی و پیش‌بینی و بهینه‌سازی مصرف انرژی را به‌طور هم‌زمان هدف قرار می‌دهد. در روش پیشنهادی، از معماری EfficientNet-B4 برای استخراج ویژگی‌های تصویری چندمقیاسی و از LSTM برای مدل‌سازی وابستگی‌های زمانی داده‌های انرژی استفاده گردید و الگوریتم چندهدف NSGA-II نیز جهت تنظیم بهینه پارامترهای فرآیندی نظیر دمای کوره، سرعت نورد و فشار به کار گرفته شد. نتایج تجربی نشان داد که مدل پیشنهادی در بخش کنترل کیفیت به‌دقت ۹۸/۷٪ و امتیاز F1 برابر با ۹۸/۵٪ دست‌یافت و عملکردی برتر نسبت به پژوهش‌های مرجع گزارش‌شده ارائه کرد. در بخش پیش‌بینی مصرف انرژی نیز مقادیر RMSE برابر با ۰/۰۸۲ و MAPE٪ برابر با ۲/۱ و مقدار R^2 برابر با ۰/۹۸ به دست آمد که بیانگر دقت بالا و پایداری مناسب مدل در مواجهه با شرایط عملیاتی مختلف است. برتری روش پیشنهادی در ترکیب هم‌زمان استخراج ویژگی‌های مکانی مبتنی بر CNN و مدل‌سازی وابستگی زمانی با LSTM نهفته است. ترکیبی که امکان شناسایی الگوهای پیچیده بین کیفیت سطح محصول و مصرف انرژی را فراهم کرده و نسبت به روش‌های تک‌بعدی عملکرد بهتری ارائه داده است. این چارچوب از منظر صنعتی می‌تواند در خطوط نورد و ریخته‌گری فولاد برای پایش بلادرنگ کیفیت سطح و ارائه پیشنهادهای خودکار جهت تنظیم بهینه فرآیند مورداستفاده قرار گیرد. چنین رویکردی علاوه بر کاهش نرخ ضایعات و دوباره‌کاری، موجب کاهش محسوس مصرف انرژی و هزینه‌های تولیدشده و با اهداف تولید سبز و اصول صنعت ٤/٠

[1] Multi-Source Learning

همسو است. بااین‌حال، پژوهش حاضر همچنان با محدودیت‌هایی همراه است. ازجمله وابستگی بخشی از داده‌های انرژی به نمونه‌های شبیه‌سازی‌شده، نیاز به ارزیابی چارچوب در خطوط تولید واقعی و شرایط غیرایستا دارد. آزمون مدل در سایر محیط‌های صنعتی و محصولات فلزی متفاوت برای سنجش تعمیم‌پذیری آن ضروری است. افزون بر این، مسیرهای متعددی برای توسعه و تکمیل این پژوهش وجود دارد. بهره‌گیری از داده‌های کاملاً واقعی و گسترده‌تر می‌تواند به ارتقای دقت و پایداری مدل کمک کند. همچنین توسعه معماری‌های چندوجهی پیشرفته‌تر مبتنی بر سازوکار توجه (Attention) می‌تواند امکان استخراج روابط پیچیده‌تر میان داده‌های سنسوری و تصویری را فراهم آورد. بررسی قابلیت‌های یادگیری انتقالی جهت کاهش نیاز به داده‌های آموزشی و درنهایت ترکیب مدل با سامانه‌های کنترل بلادرنگ و تحلیل اقتصادی–زیست‌محیطی از دیگر مسیرهای ارزشمند برای مطالعات آینده محسوب می‌شوند. با توجه به این‌که پایگاه داده NEU Surface Defect فاقد هرگونه اطلاعات زمانی یا داده‌های انرژی است، در این پژوهش از یک سازوکار هم‌تراز سازی مفهومی[1] استفاده شد. هدف از این هم‌ترازسازی، ایجاد یک چارچوب آزمایشگاهی برای بررسی امکان ترکیب داده‌های تصویری و انرژی در یک مدل واحد بوده است. در این فرآیند، برچسب‌های زمانی به‌صورت مصنوعی و صرفاً برای منطبق‌سازی نمونه‌های تصویری با داده‌های شبیه‌سازی‌شده انرژی تولید شدند. لازم به ذکر است که این هم‌ترازسازی معادل شرایط واقعی صنعتی نیست و نتایج مربوط به پیش‌بینی انرژی را باید در دسته نتایج آزمایشگاهی و مفهومی طبقه‌بندی کرد؛ بنابراین، دقت‌های گزارش‌شده در بخش انرژی مانند RMSE تنها بیانگر عملکرد مدل در یک محیط کنترل‌شده شبیه‌سازی هستند و قابل‌تعمیم مستقیم به خطوط تولید واقعی نمی‌باشند. در حال حاضر هیچ پایگاه داده صنعتی که داده‌های تصویری سطوح فولادی و داده‌های انرژی را با برچسب زمانی مشترک و واقعی ترکیب کرده باشد منتشرنشده است و پژوهش حاضر نیز مدعی ارائه چنین داده‌ای نیست. هدف اصلی از این چارچوب، ارزیابی امکان‌پذیری ترکیب چند منبعی داده‌ها و ارائه یک مدل مفهومی قابل توسعه برای کاربردهای صنعتی آتی بوده است.

References

[1] X. Zhang, Y. Chen, and S. Wang, "Surface defect detection in hot-rolled steel strips using improved YOLOv5," Sci. Rep., vol. 14, p. 1225, 2024. doi: 10.1038/s41598-024-67891-2.
[2] J. Li, Y. Wang, and Z. Liu, "Deep learning for industrial fault detection: A review," IEEE Trans. Ind. Informat., vol. 19, no. 6, pp. 5678–5690, 2023. doi: 10.1109/TII.2022.3214567.
[3] R. Kumar and A. Singh, "Machine learning models for energy consumption prediction in the steel industry," Energy Rep., vol. 9, pp. 1234–1247, 2023. doi: 10.1016/j.egyr.2023.01.045.
[4] W. Sun, Y. Zhao, and L. Ma, "Hybrid deep learning frameworks for energy-efficient manufacturing," J. Clean. Prod., vol. 453, p. 141237, 2025. doi: 10.1016/j.jclepro.2025.141237.
[5] Q. Yao, J. Chen, and Z. Li, "Integrated quality control and energy optimization in steel production using AI," J. Manuf. Syst., vol. 69, pp. 451–463, 2023. doi: 10.1016/j.jmsy.2023.07.012.
[6] H. Wang and Y. Xu, "Limitations of separate AI models for defect detection and energy prediction in industrial manufacturing," IEEE Access, vol. 12, pp. 15532–15545, 2024. doi: 10.1109/ACCESS.2024.3356712.
[7] S. Kim, J. Park, and D. Lee, "Unified deep learning framework for defect detection, predictive maintenance, and energy optimization in steel manufacturing," Appl. Energy, vol. 353, p. 121945, 2025. doi: 10.1016/j.apenergy.2025.121945.
[8] H. Liu, Y. Zhang, and M. Zhou, "Crack detection in steel surfaces using multi-scale CNN," Mater. Today Commun., vol. 36, p. 107469, 2023. doi: 10.1016/j.mtcomm.2023.107469.
[9] F. Chen, W. Li, and Q. Zhao, "IoT-based predictive maintenance system for steel manufacturing," IEEE Internet Things J., vol. 11, no. 8, pp. 13567–13578, 2024. doi: 10.1109/JIOT.2024.3345678.
[10] R. Patel, S. Kumar, and A. Mehta, "Machine vision-based defect classification in steel manufacturing," J. Manuf. Process., vol. 80, pp. 345–355, 2022. doi: 10.1016/j.jmapro.2022.07.013.

[1] Conceptual Time Alignment

[11] Z. Wang, D. Yu, and Z. Wu, "Real-time machine-learning-based optimization using input convex LSTM," arXiv, 2023. [Online]. Available: https://arxiv.org/abs/2311.07202
[12] Y. Lv, R. Hu, B. Qian, and J. B. Yang, "Q-learning-based hierarchical cooperative local search for steelmaking-continuous casting scheduling," arXiv, 2025. [Online]. Available: https://arxiv.org/abs/2506.08608
[13] N. Lee, M. Shin, A. Sagingalieva, A. J. Tripathi, K. Pinto, and A. Melnikov, "Predictive control of blast furnace temperature in steelmaking with hybrid depth-infused quantum neural networks," arXiv, 2025. [Online]. Available: https://arxiv.org/abs/2504.12389
[14] M. Abbasi, M. Plaza-Hernandez, J. Prieto, and J. M. Corchado, "Artificial intelligence of things infrastructure for quality control in cast manufacturing," Appl. Sci., vol. 15, no. 4, p. 2068, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/15/4/2068
[15] V. Varriale, A. Cammarano, F. Michelino, and M. Caputo, "Critical analysis of AI integration with cutting-edge technologies for production systems," J. Intell. Manuf., 2023. doi: 10.1007/s10845-023-02244-8.
[16] [16] D. Zhou, K. Xu, Z. Lv, J. Yang, M. Li, F. He, and G. Xu, "Intelligent manufacturing technology in the steel industry of China: A review," Sensors, vol. 22, no. 21, p. 8194, 2022. doi: 10.3390/s22218194.
[17] M. Hollander, D. A. Wolfe, and E. Chicken, Nonparametric Statistical Methods, 3rd ed. Hoboken, NJ, USA: Wiley, 2015.
[18] [18] P. R. Rosenbaum, "Exact confidence intervals for nonconstant effects by inverting the signed-rank test," J. Amer. Stat. Assoc., 2012. doi: 10.1198/0003130031405.
[19] Y. Zhang, S. Wang, Z. Jiang, J. Wang, and Y. Ma, "Strip steel surface defect detection based on lightweight YOLOv5 with feature fusion," Front. Neurorobot., vol. 17, p. 1154214, 2023. doi: 10.3389/fnbot.2023.1154214.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 141-158

Intelligent Quality Control and Defect Detection in Steel Manufacturing Processes Using Deep Learning and Image Processing for Energy Optimization

Fatemeh Saadatjou[1]*, Moslem Molalie[2]

[1]Assistant Professor, Department of Computer Engineering, Science and Arts University, Yazd, Iran

[2]Teacher, Hormozgan Education, Rudan Education Computer Group, Hormozgan, Iran

Article Information

Original Research Paper

Received:
2025 September 15

Accepted:
2025 November 17

Keywords:
Artificial Intelligence, Deep Learning, CNN-LSTM, Steel Industry

Corresponding Author*:
saadatjou@sau.ac.ir

Abstract

In recent years, the increasing competition in the steel industry and the urgent demand for reducing costs while improving quality have accelerated the adoption of artificial intelligence technologies. One of the major challenges in this domain is ensuring accurate product quality control and efficient energy management, both of which directly affect productivity and sustainability. This study proposes an integrated deep learning framework for simultaneous quality inspection and energy optimization in steel manufacturing. High-resolution surface images from the NEU surface defect dataset, combined with simulated energy consumption records, were pre-processed through filtering, normalization, and temporal alignment techniques. A hybrid multi-branch architecture, combining convolutional neural networks (CNN) for defect detection and CNN-LSTM for energy prediction, was implemented and optimized using the NSGA-II multi-objective algorithm. Experimental results demonstrate that the proposed approach achieved 98.7% accuracy and an F1-score of 98.5% in defect detection, while reducing the prediction error of energy consumption to an RMSE of 0.082 and MAPE of 2.1%. These improvements not only outperform recent state-of-the-art methods but also contribute to reducing rework, minimizing scrap rates, and achieving measurable energy savings. The proposed methodology offers a practical step toward green steel production and the realization of Industry 4.0 in heavy industries.

: 10.22034/ABMIR.2025.23661.1167

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۷۳-۱۵۹

مدل‌سازی فرآیند در کوره گندله‌سازی با استفاده از روش‌های هوشمند: مطالعه موردی در شرکت معدنی و صنعتی گل‌گهر

محدثه رضایی[۱*]، حسین نظام‌آبادی‌پور[۲]، رضا خاکسارپور[۳]

[۱] دانشجوی دکتری بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان، کرمان، ایران

[۲] استاد بخش مهندسی برق، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان، کرمان، ایران

[۳] کارشناسی ارشد، مجتمع معدنی و صنعتی گل‌گهر سیرجان، کرمان، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۴/۰۵/۳۱

تاریخ پذیرش:
۱۴۰۴/۰۸/۰۴

کلیدواژه‌ها:
مدل‌سازی هوشمند، یادگیری ماشین، انتخاب ویژگی، شبکه عصبی پرسپترون چندلایه، کوره گندله‌سازی

نویسنده‌مسئول:
mo.rezaei@eng.uk.ac.ir

چکیده

فرآیند گندله‌سازی یکی از مراحل کلیدی زنجیره تولید فولاد است که در آن کنسانتره سنگ‌آهن به گندله‌های قابل مصرف در کوره بلند تبدیل می‌شود. مرحله پخت در کوره، به‌دلیل تأثیر مستقیم توزیع دما و فشار بر ویژگی‌های فیزیکی و مکانیکی گندله، یکی از عوامل اصلی در کیفیت محصول است. هدف این پژوهش، توسعه مدلی داده‌محور برای پیش‌بینی رفتار کوره گندله‌سازی شرکت معدنی و صنعتی گل‌گهر بر پایه داده‌های عملیاتی واقعی است. بدین منظور، متغیرهای ورودی و خروجی فرآیند جمع‌آوری و پس از پیش‌پردازش، انتخاب ویژگی با سه روش امتیاز-اف، اطلاعات متقابل و ضریب همبستگی پیرسون انجام شد. سپس مدل‌های رگرسیون خطی چندگانه، جنگل تصادفی، k-نزدیک‌ترین همسایه و شبکه عصبی پرسپترون چندلایه در چارچوب چندورودی-چندخروجی برای پیش‌بینی همزمان ۳۶ متغیر دما و ۳۱ متغیر فشار آموزش داده شدند. نتایج نشان داد روش انتخاب ویژگی مبتنی بر اطلاعات متقابل موجب بهبود عملکرد مدل‌های غیرخطی شده و شبکه عصبی چندلایه بالاترین دقت را با میانگین ضریب تعیین ۹۱/۵۰ درصد برای دما و ۸۹/۶۲ درصد برای فشار به دست آورد. یافته‌ها بیانگر آن است که استفاده از رویکردهای یادگیری داده‌محور می‌تواند رفتار پیچیده کوره را با دقت بالا مدل کرده و بستر مناسبی برای طراحی سامانه‌های کنترلی هوشمند و بهینه‌سازی شرایط پخت فراهم سازد.

doi : 10.22034/ABMIR.2025.23568.1156

E-ISSN: 2821-2037

۱- مقدمه

سنگ آهن به عنوان اصلی‌ترین ماده اولیه در تولید فولاد، نقش حیاتی در صنایع معدنی و متالورژی ایفا می‌کند. این ماده معدنی حدود ۵ درصد از پوسته زمین را تشکیل می‌دهد و عمدتاً به صورت اکسیدهای هماتیت و مگنتیت یافت می‌شود [۱]. فرآیند استخراج و فرآوری سنگ آهن شامل مراحل خردایش، تغلیظ و جداسازی ناخالصی‌ها است که در نهایت به تولید کنسانتره سنگ آهن منجر می‌شود. کنسانتره، که حاوی بیش از ٦٥ درصد آهن است، برای استفاده مستقیم در کوره‌های احیای فولاد مناسب نیست، زیرا ذرات ریز آن باعث ایجاد گرد و غبار، اختلال در جریان گاز و مشکلات حمل‌ونقل می‌گردد [۲, ۳]. بنابراین، گندله‌سازی[1] به عنوان یک فرآیند کلیدی برای تبدیل کنسانتره به گلوله‌های کروی ضروری است و مواد مناسب برای احیای مستقیم یا کوره بلند را فراهم می‌کند [٤, ٥]. این فرآیند که از دهه ۱۹۵۰ میلادی آغاز شده، امروزه به یکی از گلوگاه‌های بهبود کیفیت و بهره‌وری در زنجیره تولید فولاد تبدیل شده است [٦].

در فرایند گندله‌سازی، کنسانتره با مواد افزودنی مانند بنتونیت، سنگ آهک و آب مخلوط شده و در دیسک‌ها یا استوانه‌های چرخان به گلوله تبدیل می‌شود، سپس در کوره پخت حرارت می‌بیند تا استحکام مکانیکی لازم را کسب کند [۷]. روش‌های اصلی گندله‌سازی شامل شبکه-کوره-خنک‌کن[2] و شبکه مستقیم[3] است که هر کدام مزایا و چالش‌های خاصی دارند [۸]. برای مثال، روش شبکه مستقیم که در شرکت معدنی و صنعتی گل‌گهر مورد استفاده قرار می‌گیرد، با کاهش خرد شدن گندله‌ها و تنظیم دقیق عملیات حرارتی، کیفیت بالاتری ارائه می‌دهد [۹].

کوره گندله‌سازی مورد مطالعه در این پژوهش، از نوع شبکه مستقیم و با طول ۱۵٦ متر، شامل نواحی خشک‌کن دمشی[4] (UDD) و مکشی[5] (DDD)، پیش‌گرمایش[6] (PH)، پخت[7] (F و AF) و خنک‌کن[8] (C1 و C2) است [۱۰]. در این فرآیند، دما از شرایط محیطی به حدود ۱۳۰۰ درجه سانتی‌گراد در مرحله پخت افزایش یافته و سپس در نواحی خنک‌کن تا حدود ۲۰۰ درجه سانتی‌گراد کاهش می‌یابد. کیفیت گندله‌های تولید شده، به شدت تحت تأثیر نسبت ترکیب مواد خام، ارتفاع بستر و شرایط حرارتی نظیر دما و فشار در بخش‌های مختلف کوره قرار دارد. تغییرات ناخواسته در این پارامترها می‌تواند منجر به کاهش استحکام مکانیکی، افزایش مصرف انرژی و بروز مشکلات زیست‌محیطی شود [۱۱].

در عمل، کنترل دما و فشار اغلب بر تجربه اپراتورها متکی است که دست‌یابی به تعادل حرارتی پایدار را دشوار می‌سازد. از این‌رو، استفاده از مدل‌های پیش‌بینی مبتنی بر داده، برای ارزیابی اثر تغییرات ورودی پیش از اجرای واقعی، ضروری به نظر می‌رسد. در سال‌های اخیر، مطالعات متعددی از هوش مصنوعی و یادگیری ماشین برای مدل‌سازی و کنترل فرآیندهای پیچیده صنعتی ازجمله گندله‌سازی استفاده کرده‌اند. برای مثال، مدل‌های یادگیری عمیق در پیش‌بینی اندازه گندله، کنترل فرآیند پخت و بهینه‌سازی مصرف انرژی به‌کار رفته‌اند [۱۲]، و شبکه‌های عصبی پیچشی[9] (CNN) نیز برای پایش برخط اندازه گندله توسعه یافته‌اند [۱۳]. نتایج این پژوهش‌ها نشان داده که استفاده از فناوری‌های یادگیری ماشین می‌تواند دقت پیش‌بینی و کارایی فرآیندهای معدنی همچون تف‌جوشی و گندله‌سازی را افزایش داده و هم‌زمان با بهبود بهره‌وری انرژی، به پایداری زیست‌محیطی کمک کند [۱٤-۱٦].

در حوزه مدل‌سازی فرآیندهای حرارتی، دو رویکرد اصلی در ادبیات علمی وجود دارد: مدل‌های مکانیکی/ فیزیکی[10] و مدل‌های داده‌محور[11]. مدل‌های فیزیکی بر مبنای قوانین انتقال حرارت، جرم و واکنش‌های سینتیکی توسعه یافته‌اند و می‌توانند رفتار فرآیند را با دقت بالا توصیف کنند. برای نمونه، در [۸]، مدلی ریاضی برای فرآیند استحکام گندله‌ها بر پایه شبکه مستقیم ارائه شده که توزیع دما، نرخ خشک شدن و واکنش‌های اکسیداسیون را شبیه‌سازی

[1] Pelletizing
[2] Grate-Kiln Cooler
[3] Straight-Grate
[4] Updraft Drying
[5] Downdraft Drying
[6] Pre-Heating
[7] Firing
[8] Cooling
[9] Convolutional Neural Network (CNN)
[10] Mechanistic/Physical Models
[11] Data-Driven Models

می‌کند. همچنین در [۱۷, ۱۸]، مدل‌های فیزیکی جامعی برای تحلیل واکنش‌های شیمیایی، اکسیداسیون مگنتیت و احتراق سوخت در نواحی مختلف کوره ارائه شده است. با وجود دقت بالا، این مدل‌ها نیازمند داده‌های آزمایشگاهی دقیق و توان محاسباتی زیاد هستند، ازاین‌رو در کاربردهای کنترل برخط چندان عملیاتی نیستند.

در مقابل، مدل‌های داده‌محور با استفاده از داده‌های واقعی کارخانه و الگوریتم‌های یادگیری ماشین قادرند رفتار سیستم را بدون نیاز به مدل‌سازی دقیق فیزیکی پیش‌بینی کنند. مطالعات متعددی [۱۹, ۲۰] نشان داده‌اند که رویکردهای یادگیری ماشین، از جمله شبکه‌های عصبی، رگرسیون جنگل تصادفی و روش‌های ترکیبی، می‌توانند توزیع دما در انواع کوره‌ها را با دقت بالاتری نسبت به مدل‌های ساده فیزیکی پیش‌بینی کنند و در کاربردهای بلادرنگ نیز قابل پیاده‌سازی‌اند.

بر این اساس، پژوهش حاضر بر استخراج مدل ورودی–خروجی برای کوره گندله‌سازی با استفاده از روش‌های هوش مصنوعی تمرکز دارد. اهداف شامل جمع‌آوری و پیش‌پردازش داده‌ها، تحلیل آماری روابط متغیرها، اعمال شیفت زمانی برای همگام‌سازی ورودی–خروجی، انتخاب ویژگی و مدل‌سازی با الگوریتم‌های یادگیری ماشین مانند رگرسیون خطی چندگانه[1] (MLR)، رگرسیون جنگل تصادفی[2] (RFR)، رگرسیون k–همسایه نزدیک‌تر[3] (KNR) و شبکه‌های عصبی عمیق است. این رویکرد می‌تواند ابزار مفیدی برای کنترل پیشرفته فرآیند و دستیابی به گندله با کیفیت مطلوب فراهم کند.

نوآوری‌های اصلی این پژوهش را می‌توان در موارد زیر خلاصه نمود:

- استفاده از یک مجموعه داده یک ساله با گام زمانی یک دقیقه از کوره گندله‌سازی صنعتی به‌عنوان داده واقعی (۵۲۵،۵۴۰ نمونه) برای آموزش و اعتبارسنجی مدل‌ها.
- مقایسه نظام‌مند سه روش انتخاب ویژگی تک‌متغیره انتخاب–اف، اطلاعات متقابل و همبستگی پیرسون در ترکیب با سه الگوریتم رگرسیون (MLR, KNR, RFR) بر روی ۶۷ خروجی عملیاتی.
- طرح و ارزیابی مدل‌های چندورودی–چندخروجی[4] (MIMO) برای پیش‌بینی هم‌زمان دما و فشار در نواحی مختلف کوره.
- ارائه تحلیل عملیاتی (پیشنهاد برای کاربرد در سیستم‌های کنترلی صنعتی) و راهنمایی برای انتخاب مدل بر اساس محدودیت‌های محاسباتی و مقیاس خروجی.

ادامه مقاله بدین شرح سازمان‌دهی شده است: در بخش دوم مرور ادبیات موضوعی و در بخش سوم مفاهیم پایه‌ای که برای هدف این پژوهش مورد توجه هستند، بیان می‌شوند. در بخش چهارم روش پیشنهادی و در بخش پنجم آزمایش‌ها و نتایج بدست آمده ارائه می‌شود. در بخش ششم نتیجه‌گیری و پیشنهادات برای تحقیقات آینده بیان می‌شود.

۲– مفاهیم پایه

برای درک بهتر روش پیشنهادی و تحلیل نتایج، آشنایی با مفاهیم اصلی صنعتی و علمی مرتبط با موضوع پژوهش ضروری است. در این بخش، ابتدا مروری بر فرآیند گندله‌سازی و عملکرد کوره پخت ارائه می‌شود و سپس مفاهیم یادگیری ماشین و شبکه عصبی پرسپترون چندلایه که در مدل‌سازی به کار رفته‌اند، توضیح داده خواهد شد.

۲–۱ فرآیند گندله‌سازی

گندله‌سازی فرآیندی است که در آن کنسانتره سنگ‌آهن با افزودنی‌هایی مانند بنتونیت، سنگ آهک و آب مخلوط شده و به شکل گلوله‌های کروی با قطر ۸ تا ۱۶ میلی‌متر در می‌آید. هدف از این عملیات، بهبود قابلیت حمل، کاهش گرد و غبار، و فراهم‌سازی شرایط بهینه برای فرآیندهای بعدی مانند احیای مستقیم یا استفاده در کوره بلند است.

پس از تشکیل گندله خام، این ذرات وارد کوره پخت می‌شوند تا با اعمال پروفیل حرارتی مناسب، استحکام مکانیکی و ویژگی‌های فیزیکی مورد نظر را به‌دست آورند. کوره‌های گندله‌سازی به دو دسته اصلی تقسیم می‌شوند: الف) روش شبکه–کوره–خنک‌کن، ب) روش شبکه مستقیم. در کوره‌های شبکه مستقیم، از جمله کوره مورد مطالعه در این پژوهش، فرآیند پخت در چند ناحیه متوالی

[1] Multiple Linear Regression
[2] Random Forest Regression
[3] k-Nearest Neighbor Regression
[4] Multiple Input Multiple Output

انجام می‌شود. شکل (۱) طرح کلی کوره گندله‌سازی را نشان می دهد.

- خشک‌کن دمشی (UDD) و مکشی (UDD): تبخیر رطوبت گندله‌ها با جریان هوای گرم از پایین یا بالا. دمای خروجی این دو ناحیه حدود ۳۵۰ درجه سانتی‌گراد است.
- پیش‌گرمایش (PH): افزایش تدریجی دما از ۳۵۰ درجه سانتی‌گراد تا حدود ۱۰۰۰ درجه سانتی‌گراد.
- پخت (F و AF): رسیدن دما به حدود ۱۳۰۰ درجه سانتی‌گراد برای ایجاد استحکام نهایی. در AF با حفظ گرما و بدون مشعل، دما ثابت می‌ماند.
- خنک‌کن (C1 و C2):کاهش دما تا حدود ۲۰۰ درجه سانتی‌گراد همراه با بازیافت حرارت و بازگرداندن هوای گرم به نواحی پخت و خشک‌کن، با هدف بهینه‌سازی مصرف انرژی.

در انتهای این فرآیند، گندله‌هایی با ویژگی‌های مکانیکی و شیمیایی مطلوب برای مراحل بعدی تولید فولاد حاصل می‌شوند. شکل (۲) نمایی از کارخانه گندله‌سازی شماره ۱ شرکت معدنی و صنعتی گل‌گهر را نشان می‌دهد که داده‌های مورد استفاده در این پژوهش از کوره گندله‌سازی همین واحد استخراج شده‌اند.

۲-۲ الگوریتم‌های یادگیری ماشین

در این بخش، به معرفی مفاهیم نظری الگوریتم‌های مورد استفاده برای انتخاب ویژگی و مدل‌سازی داده‌های فرآیند صنعتی در این پژوهش پرداخته می‌شود. این مدل‌ها شامل روش‌های کلاسیک یادگیری ماشین و شبکه‌های عصبی مصنوعی هستند که هر یک دارای کاربردها و مزایای خاص خود در تحلیل داده‌های چندمتغیره صنعتی می‌باشند.

در این پژوهش برای انتخاب ویژگی از سه روش فیلتر آماری و برای مدل‌سازی داده‌های فرآیند صنعتی، از روش‌های مختلف یادگیری ماشین استفاده شده است. در مرحله انتخاب ویژگی، شاخص‌هایی مانند امتیاز-اف [۲۲]، اطلاعات متقابل [۲۳] و ضریب همبستگی پیرسون به‌کار رفتند تا متغیرهای مؤثر شناسایی شوند. برای مدل‌سازی، چند رویکرد مورد استفاده قرار گرفت:

- رگرسیون خطی چندگانه به عنوان روشی ساده برای تحلیل روابط خطی [۲۴].
- جنگل تصادفی با قابلیت بالای مدل‌سازی روابط غیرخطی و تعیین اهمیت ویژگی‌ها [۲۵].
- رگرسیون k-همسایه نزدیک‌تر به عنوان روشی ساده ولی کارآمد در پیش‌بینی [۲۶].
- شبکه‌های عصبی مصنوعی برای یادگیری الگوهای پیچیده و غیرخطی داده‌ها [۲۷, ۲۸].

این ترکیب از روش‌ها امکان تحلیل داده‌های چندمتغیره صنعتی را از جنبه‌های مختلف فراهم ساخته و موجب افزایش دقت و اطمینان نتایج می‌شود.

۳- روش پیشنهادی

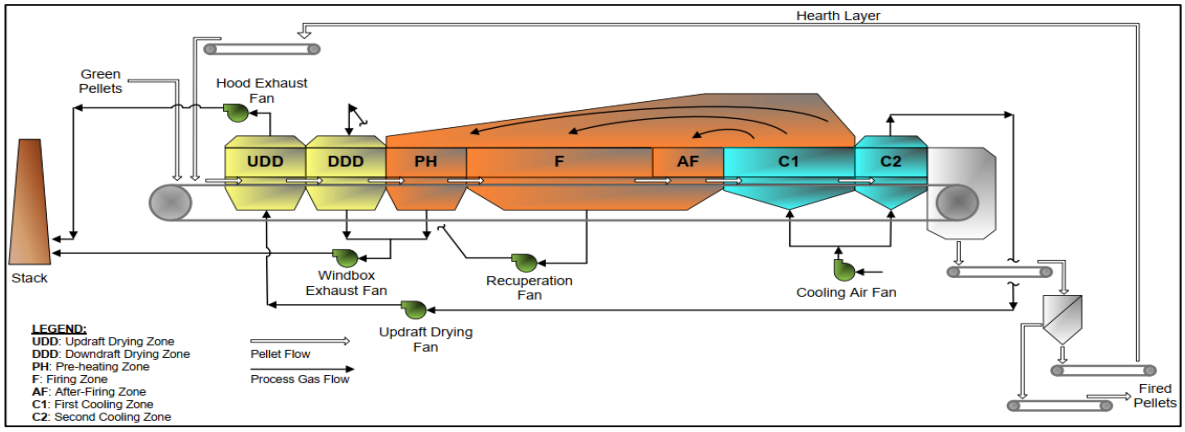
با توجه به ماهیت پیچیده، غیرخطی و چندمتغیره‌ی فرآیند گندله‌سازی، استفاده از مدل‌های تحلیلی مبتنی بر قوانین فیزیکی به‌سختی امکان‌پذیر است. از این‌رو، در این پژوهش یک رویکرد داده‌محور نظارت‌شده مبتنی بر الگوریتم‌های یادگیری ماشین و یادگیری عمیق به‌منظور مدل‌سازی رفتار فرآیند و پیش‌بینی هم‌زمان فشار و دما در نواحی مختلف کوره پیشنهاد می‌شود. این روش به دلیل توانایی در یادگیری روابط غیرخطی و بهره‌گیری از داده‌های واقعی صنعتی برای مسائل پیچیده‌ی فرآیندی، مناسب‌ترین گزینه برای هدف این تحقیق است.

در مرحله نخست، کلان‌داده (۵۲۵،۵۴۰ نمونه) حاصل از سامانه پایش کارخانه گندله‌سازی شماره ۱ شرکت معدنی و صنعتی گل‌گهر گردآوری و پردازش شد. داده‌ها به دو گروه ورودی و خروجی تقسیم شدند، مجموعه‌ی ورودی شامل ٦٦ متغیر فرآیندی (۳٤ متغیر دمای مشعل‌های نواحی PH و F، ۲۳ متغیر درصد دمپر هواکش‌ها، ٤ متغیر خوراک ورودی، یک متغیر سرعت ماشین، سه متغیر ارتفاع بستر گندله و یک متغیر میانگین ارتفاع بستر) و مجموعه‌ی خروجی شامل ٦٧ متغیر (۳۱ فشار و ۳٦ دما در نواحی مختلف کوره) می‌باشد.

در فرآیند نگاشت پارامتری، ویژگی‌هایی مانند دمای مشعل‌ها، درصد بازبودن دمپرها، سرعت نوار نقاله، ارتفاع بستر و نرخ خوراک ورودی، به‌عنوان بردار ورودی مدل ($\boldsymbol{x} = [x_1, x_2, \dots, x_n]$) در نظر گرفته شدند که n بیانگر تعداد متغیرهای ورودی است. در مقابل، بردار خروجی شامل فشار و دما در نواحی

مختلف کوره ($\boldsymbol{y} = [y_1, y_2, \dots, y_m]$) است که m نشان‌دهنده‌ی تعداد متغیرهای خروجی است. هر ردیف از داده‌ها معرف یک نمونه از اندازه‌گیری‌های فرآیند در زمان معین بوده و در مجموع شامل N نمونه آموزشی است. بدین ترتیب، مدل‌های یادگیری ماشین و شبکه‌ی عصبی پرسپترون چندلایه به منظور یادگیری نگاشت غیرخطی بین ورودی‌ها و خروجی‌ها ($y = f(x)$) آموزش داده شدند. این چارچوب، امکان مدل‌سازی هم‌زمان اثرات متقابل متغیرهای فرآیندی و پیش‌بینی پاسخ‌های چندگانه را فراهم می‌کند.

روش پیشنهادی شامل دو مسیر مکمل است: مسیر (الف) برای انتخاب ویژگی‌های مؤثر با استفاده از سه الگوریتم فیلترینگ آماری، و مسیر (ب) مدل‌سازی رابطه‌ی بین ورودی‌ها و خروجی‌ها با بهره‌گیری از رویکردهای یادگیری ماشین و یادگیری عمیق. این ساختار به‌گونه‌ای طراحی شده که علاوه بر افزایش دقت پیش‌بینی، تفسیرپذیری و کارایی محاسباتی مدل نیز حفظ شود.



شکل (۱): نمای کلی کوره گندله‌سازی [۱]

شکل (۲): کارخانه گندله‌سازی شماره ۱ شرکت معدنی و صنعتی گل‌گهر.

۳-۱ انتخاب ویژگی فیلتری

همان‌طور که گفته شد، در مسیر الف، انتخاب ویژگی با روش‌های فیلتری معرفی شده در بخش ۲-۲ انجام می‌شود. به ازای هر متغیر خروجی، ده ویژگی برتر هر معیار برای هر متغیر خروجی انتخاب شدند و جدول ۱ نتایج این انتخاب را برای متغیرهای خروجی فشار نواحی UDD و DDD که به ترتیب با نمادهای UDD_P و DDD_P نشان داده می‌شوند، نشان می‌دهد.

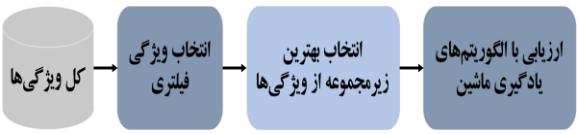
بر اساس داده‌های این جدول، برای متغیر خروجی UDD_P ، مشاهده می‌شود که از میان ۱۰ ویژگی برتر که توسط هر الگوریتم انتخاب شده‌اند، فقط یک ویژگی در هر سه فهرست مشترک است.

یعنی دو الگوریتم دیگر روی انتخاب‌های متفاوتی تمرکز کرده‌اند و تنها یک متغیر به‌طور همزمان توسط هر سه روش مهم تشخیص داده شده است. اما در مورد متغیر خروجی DDD_P ، سه الگوریتم انتخاب ویژگی، ۸ ویژگی مشترک و تنها ۲ ویژگی متفاوت را انتخاب کرده‌اند. همان‌گونه که مشاهده می‌شود، تفاوت میان ویژگی‌های منتخب در سه معیار نشان‌دهنده‌ی دیدگاه‌های مکمل آن‌ها است. به‌ویژه، معیار اطلاعات متقابل متغیرهایی را آشکار ساخته که علی‌رغم نداشتن همبستگی خطی قوی، در رفتار کلی فرآیند اثرگذار هستند. استفاده‌ی همزمان از این سه روش، تصویری جامع‌تر از ساختار داده‌ها فراهم کرده و مبنایی مناسب برای مدل‌سازی دقیق‌تر فراهم ساخته است.

جدول (۱): ۱۰ ویژگی منتخب توسط هر یک از سه روش انتخاب ویژگی برای متغیرهای خروجی UDD_P و DDD_P

ویژگی	UDD_P			DDD_P		
	انتخاب-اف	اطلاعات متقابل	همبستگی پیرسون	انتخاب-اف	اطلاعات متقابل	همبستگی پیرسون
۱	DDDtoHEX_D	DDDtoHEX_D	DDDtoHEX_D	DDDtoHEX_D	DDDtoHEX_D	DDDtoHEX_D
۲	UDDtoHEX_D	Oct_D	UDDtoHEX_D	F36_D	F22_D	F36_D
۳	B3_T	F22_D	B14_T	F37_D	F37_D	F37_D
٤	F22_D	F37_D	B12_T	UDDtoHEX_D	F36_D	UDDtoHEX_D
٥	F12_D	F36_D	B27_T	GPtoIM_FR	F42_D	GPtoIM_FR
٦	B6_T	F42_D	Oct27B_D	GPtoRS_FR	IM_S	GPtoRS_FR
۷	B5_T	F22toF12_D	Oct27A_D	Num_Disks	F002_D	Num_Disks
۸	B14_T	B3_T	F22toF12_D	F42_D	B33_T	F42_D
۹	B4_T	B6_T	GPtoIM_FR	IM_S	Oct_D	IM_S
۱۰	B2_T	B19_T	B11_T	B27_T	B31_T	B27_T

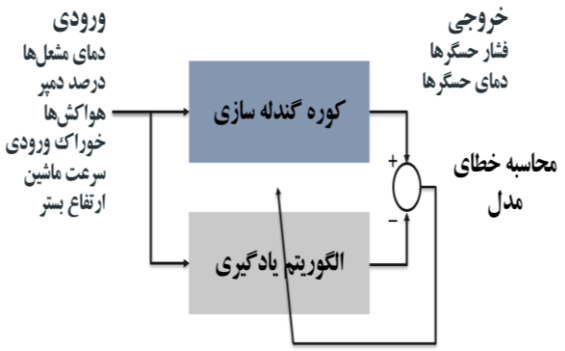
این مرحله برای ٦۷ متغیر خروجی موجود در مجموعه‌داده انجام شده و ۱۰ ویژگی برتر برای هر یک از آن‌ها شناسایی شده‌اند. از آنجا که برای ارزیابی مستقیم کیفیت الگوریتم‌های انتخاب ویژگی فیلتری معیار مشخص و مستقلی وجود ندارد، عملکرد آن‌ها به‌طور غیرمستقیم و از طریق تأثیرشان بر نتایج مدل‌های رگرسیونی سنجیده می‌شود. بدین منظور، از سه مدل MLR، KNR و RFR استفاده شده است. نمودار بلوکی این مرحله در شکل (۳) نشان داده شده است.



شکل (۳): نمودار بلوکی مسیر (الف)

۳-۲ مدل‌سازی

در بخش ۳-۱، مدل‌سازی با روش‌های سنتی یادگیری ماشین بر مبنای ویژگی‌های منتخب توسط سه الگوریتم انتخاب ویژگی انجام شد تا اثر نوع روش انتخاب ویژگی بر عملکرد مدل‌های رگرسیون بررسی شود. در این بخش، با هدف مقایسه و ارزیابی جامع‌تر، مدل‌سازی مجدداً انجام می‌شود. اما این بار همه متغیرهای ورودی بدون اعمال فیلتر انتخاب ویژگی در نظر گرفته شده‌اند. این کار با مدل‌های تک خروجی MLR، KNR و RFR و نیز یک مدل مبتنی بر MLP به نام MIMO انجام می‌شود. نمودار بلوکی مسیر (ب)، یعنی فرآیند مدل‌سازی داده‌محور در شکل ٤ نمایش داده شده است. در این چارچوب، متغیرهای ورودی وارد مدل یادگیری می‌شوند تا رابطه‌ی غیرخطی میان ورودی‌ها و خروجی‌های فرآیند یاد گرفته شود.



شکل (٤): نمودار بلوکی مسیر (ب)

همان‌طور که در شکل مشاهده می‌شود، مدل یادگیری به‌صورت یک حلقه بازخوردی عمل می‌کند. به‌گونه‌ای که خروجی پیش‌بینی‌شده‌ی مدل با مقادیر واقعی اندازه‌گیری‌شده مقایسه و خطای مدل محاسبه می‌شود. سپس الگوریتم یادگیری با استفاده از این خطا، پارامترهای مدل را به‌روزرسانی می‌کند تا همگرایی به نگاشت بهینه $Y = f(X)$ حاصل شود. این ساختار امکان بهبود تدریجی دقت پیش‌بینی و مدل‌سازی هم‌زمان اثرات متقابل میان متغیرهای فرآیندی را فراهم می‌کند.

در مدل‌های سنتی مانند MLR، KNR و RFR، به دلیل ساختار تک‌خروجی آن‌ها، لازم است برای پیش‌بینی هر یک از متغیرهای خروجی، یک مدل جداگانه آموزش داده شود. این امر موجب افزایش تعداد مدل‌های مورد نیاز، زمان آموزش و پیچیدگی محاسباتی می‌گردد. در مقابل، مدل MIMO قادر است به‌طور همزمان تمام خروجی‌های فرآیند را پیش‌بینی کند. این ویژگی باعث کاهش تعداد مدل‌ها، بهره‌برداری بهتر از روابط متقابل بین خروجی‌ها و در بسیاری از موارد بهبود دقت پیش‌بینی می‌شود .

مدل MIMO به صورت یک شبکه عصبی پرسپترون چندلایه طراحی شده است. این مدل‌سازی به سه صورت انجام شده است:

MIMO-Temperature: آموزش یک مدل بر اساس تمام متغیرهای ورودی و پیش‌بینی همزمان همه خروجی‌های دما.

MIMO- Pressure: آموزش یک مدل بر اساس تمام متغیرهای ورودی و پیش‌بینی همزمان همه خروجی‌های فشار.

MIMO-All: آموزش یک مدل بر اساس تمام متغیرهای ورودی و پیش‌بینی همزمان همه خروجی‌های فشار و دما.

٤- آزمایش‌ها و نتایج

در این بخش، فرآیند ارزیابی مدل‌های پیشنهادی و تحلیل نتایج ارائه می‌شود. ابتدا تنظیمات آزمایش‌ها، شامل پیکربندی و پیش‌پردازش داده‌ها، جداسازی داده‌های آموزش و آزمون، معیارهای ارزیابی عملکرد و مقادیر پارامترهای هر الگوریتم برای هر مدل توضیح داده می‌شود. سپس عملکرد مدل‌ها مورد بررسی قرار می‌گیرد.

مقایسه بین مدل‌های سنتی (MLR، KNR، RFR) و مدل MIMO در ساختارهای MIMO_Temperature، MIMO_Pressure و MIMO_All انجام می‌شود. در ادامه، نتایج بدست آمده تحلیل و مقایسه می‌گردند تا نقاط قوت و ضعف هر رویکرد مشخص شود.

٤-١ معیارهای ارزیابی عملکرد

برای سنجش دقت مدل‌های رگرسیونی، از سه معیار متداول استفاده شده است:

- میانگین مربعات خطا (MSE): نشان‌دهنده میانگین توان دوم اختلاف بین مقادیر واقعی و پیش‌بینی شده بوده و مقادیر کمتر بیانگر عملکرد بهتر مدل است [۲۹].
- میانگین قدر مطلق خطا (MAE): مشابه MSE، اما بر پایه قدرمطلق خطا محاسبه می‌شود و معیار ساده‌تری برای ارزیابی دقت پیش‌بینی است [۳۰].
- ضریب تعیین: که با نماد R^2-Score نشان داده می‌شود، میزان تغییرات توضیح داده شده توسط مدل را نشان می‌دهد و بین صفر تا یک متغیر است. مقدار بالاتر نشان‌دهنده دقت و توانایی بالاتر مدل در تبیین داده‌هاست [۱۰].

در مسأله حاضر، از دو معیار MSE و MAE برای ارزیابی عملکرد مدل‌ها استفاده شده است. زیرا هر یک جنبه متفاوتی از خطا را نشان می‌دهند. معیار MSE به دلیل توان دوم کردن خطاها، نسبت به خطاهای بزرگ حساس‌تر است و برای تحلیل مدل‌هایی که هدف آن‌ها کاهش خطاهای شدید و پایدارسازی پیش‌بینی‌هاست (مانند مدل‌های مبتنی بر شبکه عصبی و جنگل تصادفی)، شاخص مناسب‌تری محسوب می‌شود. در مقابل، معیار MAE میانگین قدرمطلق خطاها را بدون تأکید بر مقدار آن‌ها ارائه می‌دهد و در سنجش دقت کلی پیش‌بینی‌ها (به‌ویژه برای متغیرهایی با مقیاس‌های مختلف) قابل تفسیرتر است.

٤-٢ پیش‌پردازش داده‌ها

داده‌های موجود، شامل داده‌های یک ساله کوره گندله‌سازی هستند که از تاریخ اول ژانویه ۲۰۲۱ الی ۳۱ دسامبر ۲۰۲۱ با گام زمانی یک دقیقه جمع‌آوری شده‌اند. در این پژوهش، ٦٦ متغیر به عنوان ویژگی‌های ورودی در نظر گرفته شده‌اند که شامل دمای مشعل‌ها، درصد دمپر هواکش‌ها، سرعت ماشین، ارتفاع بستر گندله و میزان خوراک ورودی هستند. در مقابل، ٦٧ متغیر خروجی شامل مقادیر اندازه‌گیری‌شده‌ی دما و فشار در نقاط مختلف کوره می‌باشند که

مدل باید آن‌ها را بر اساس ورودی‌ها پیش‌بینی کند.

در مجموع، مجموعه‌داده شامل ۱۳۳ ویژگی (٦٦ ورودی و ٦٧ خروجی) و ۵۲۵٬۵۴۰ نمونه است. این داده‌ها پس از پاک‌سازی و پیش‌پردازش برای آموزش مدل‌های یادگیری ماشین استفاده شده‌اند. مراحل پیش‌پردازش مجموعه داده شامل مواجهه با مقادیر ناموجود، مقیاس‌بندی ویژگی با روش نرمال‌سازی کمینه–بیشینه در بازه [−۱،۱] و تقسیم به داده‌های آموزش و آزمون است.

٤-٢-١ مواجهه با مقادیر ناموجود

برخی از داده‌ها دارای مقادیر ناموجود (NaN) هستند که لازم است پیش از تحلیل تکمیل شوند. این مقادیر عمدتاً در شرایطی مشاهده شده‌اند که تعداد دیسک‌های فعال کمتر از سه بوده است. جدول (۲) میانگین تعداد مقادیر ناموجود را برای داده‌های ورودی و خروجی ارائه می‌کند.

جدول (۲): میانگین تعداد مقادیر ناموجود به ازای هر متغیر مجموعه داده

نام مجموعه داده	میانگین تعداد مقادیر ناموجود در هر متغیر
داده‌های دمای مشعل	۵۸۳۳٦
داده‌های درصد دمپر هواکش‌ها	۵۸۳۲۸
داده‌های میزان خوراک ورودی	۵۸۹۹۹
داده‌های سرعت ماشین و ارتفاع بستر	۵۹۰۹۳
داده‌های فشار حسگرها	۵۸۹۸۱
داده‌های دمای حسگرها	۵۸٤۲٤

به عنوان نمونه، سطر اول این جدول به داده‌های دمای مشعل‌ها مربوط است. این داده‌ها شامل ۳٤ متغیر دماست که به طور میانگین، هر متغیر دارای ۵۸۳۳٦ مقدار ناموجود است. پیش از انجام تحلیل‌های آماری، این مقادیر باید تکمیل گردند. با توجه به اینکه تغییرات این متغیرها در طول زمان چندان زیاد نیست، در این پژوهش از روش جایگزینی مقدار میانگین هر متغیر برای پر کردن مقادیر ناموجود آن استفاده شده است.

٤-٢-٢ مقیاس‌بندی ویژگی

پیش از به‌کارگیری هر الگوریتم یادگیری ماشین، معمولاً مرحله‌ای تحت عنوان مقیاس‌بندی ویژگی‌ها بر روی داده‌ها انجام می‌شود. هدف از این مرحله، همسان‌سازی دامنه مقادیر متغیرهای مختلف است تا الگوریتم بتواند آن‌ها را به شکلی متعادل پردازش کند. از این رو، پیش از اجرای الگوریتم‌های انتخاب ویژگی نیز لازم است داده‌ها مقیاس‌بندی شوند. در روش نرمال‌سازی کمینه–بیشینه، به منظور تغییر مقیاس خطی ویژگی j–ام در یک محدوده دلخواه [a,b] می‌توان از رابطه (۱) استفاده کرد [۳۱]:

$$\hat{x}_j^i = a + \frac{(x_j^i - x_j^{min})(b-a)}{x_j^{max} - x_j^{min}} \qquad (۱)$$

در رابطه بالا، x_j^{min} و x_j^{max} به ترتیب برابر با کمینه و بیشینه مقادیر ویژگی j–ام و $\hat{x}_j^i$ مقدار مقیاس شده را نشان می‌دهد. در این پژوهش از نرمال‌سازی کمینه–بیشینه استفاده شده است و کلیه مقادیر متغیرهای ورودی و خروجی به بازه $[-1,1]$ نگاشت شده‌اند.

لازم به ذکر است که نرمال‌سازی فقط برای آموزش مدل انجام شد. برای گزارش عملکرد، مقادیر پیش‌بینی شده توسط مدل، به مقیاس اصلی بازگردانده شده و سپس معیارهای ارزیابی عملکرد مانند MAE و MSE محاسبه شده‌اند.

٤-٢-٣ جداسازی داده‌های آموزش و آزمون

در این پژوهش، هدف مدل‌سازی، پیش‌بینی همزمان خروجی‌های فرآیند بر اساس متغیرهای ورودی در همان لحظه است، نه پیش‌بینی روند زمانی آینده . در چنین مسائلی، مدل به‌دنبال کشف رابطه میان مقادیر ورودی و خروجی در هر نمونه از داده است، بدون آن‌که به توالی زمانی داده‌ها وابسته باشد. از آنجا که ورودی‌ها و خروجی‌ها مربوط به اندازه‌گیری‌های لحظه‌ای از وضعیت فرآیند هستند و در بازه‌های زمانی نسبتاً کوتاه ثبت شده‌اند، ترتیب وقوع داده‌ها در آموزش مدل اهمیتی ندارد و بر روابط فیزیکی میان متغیرها تأثیر نمی‌گذارد.

به همین دلیل، تقسیم داده‌ها به‌صورت تصادفی میان مجموعه‌های آموزش (۷۰ درصد) و آزمون (۳۰ درصد) انجام شد تا هر دو مجموعه شامل گستره کامل شرایط عملیاتی باشند. الگوریتم‌های مورد مقایسه با استفاده از داده‌های آموزشی آموزش دیده و سپس عملکرد آن‌ها بر روی داده‌های آزمون ارزیابی می‌شود.

۴-۳ تنظیم پارامترها

برای مرحله انتخاب ویژگی از ماژول‌های موجود در کتابخانه Scikit-learn پایتون استفاده شده است. به کمک ماژول SelectKBest این کتابخانه و انتخاب تابع رتبه‌دهی با امتیاز-اف، اطلاعات متقابل و همبستگی پیرسون، رتبه¬های مربوط به هر ویژگی محاسبه شده و بر این اساس، به ازای هر متغیر خروجی کوره گندله‌سازی، ۱۰ ویژگی ورودی با بالاترین رتبه-ها انتخاب می‌شوند.

در ادامه لازم است مدل‌های رگرسیون MLR، KNR و RFR روی ویژگی‌های انتخاب شده برازش داده شده و مقادیر متغیرهای خروجی را پیش‌بینی کنند. این مدل‌ها با استفاده از کتابخانه cuML (بخشی از چارچوب RAPIDS) پیاده‌سازی شده‌اند که امکان اجرای الگوریتم‌ها بر بستر GPU را فراهم می‌کند. استفاده از این کتابخانه با هدف افزایش سرعت پردازش و مدیریت کارآمد حجم بالای داده‌ها انتخاب شده است. تنظیمات پارامترها برای الگوریتم‌های MLR، KNR، RFR در جدول (۳) ارائه شده است.

جدول (۳): تنظیمات پارامترهای مدل‌های MLR، KNR و RFR

الگوریتم	پارامترها
MLR	$algorithm = 'eig', fit\_intercept = True, normalize = False$
KNR	$n\_neighbors = 5,$ $metric = 'euclidean', weights = 'uniform',$
RFR	$n\_estimator = 100, split\_criterion = 'mse', max\_depth = 16$

مدل MIMO، با استفاده از لایه‌های پیش‌خور طراحی شده است. از آنجا که داده‌های ورودی شامل متغیرهای عددی مستقل و بدون ساختار مکانی هستند، استفاده از مدل‌های کانولوشنی در این مسئله توجیه فیزیکی و محاسباتی ندارد زیرا بین ویژگی‌ها رابطه مکانی یا همسایگی تعریف‌پذیر وجود ندارد. بنابراین از ساختار تماماً متصل برای همه لایه‌ها استفاده شده است. این شبکه دارای یک لایه ورودی با ٦٦ گره برای دریافت متغیرهای ورودی است. پس از لایه ورودی، ما از پنج لایه پنهان که به ترتیب شامل ۱۳۲، ۲٦٤، ۲٦٤، ۲٦٤، ۱۳۲ گره هستند استفاده کرده‌ایم. لایه خروجی در مدل‌های MIMO_All، MIMO_Pressure و MIMO_Temperature به ترتیب شامل ٦٧ (تعداد کل متغیرهای خروجی)، ۳۱ (تعداد متغیرهای فشار) و ۳٦ (تعداد متغیرهای دما) گره است.

انتخاب ساختار پنج‌لایه با هدف افزایش توان مدل در یادگیری روابط غیرخطی میان ورودی‌ها و خروجی‌ها صورت گرفت. پیش از انتخاب این معماری، نسخه‌های ساده‌تر شبکه با دو و سه لایه پنهان نیز آزمایش شدند. اما با وجود زمان آموزش کوتاه‌تر، این مدل‌ها دقت پایین‌تر و نوسان بیشتری در خطای اعتبارسنجی داشتند. بنابراین ساختار پنج‌لایه به عنوان پیکربندی نهایی برگزیده شد.

تابع فعال‌ساز در لایه‌های پنهان ReLU و در لایه خروجی، خطی است. تابع زیان مورد استفاده، MSE بوده و بهینه¬ساز ADAM با نرخ یادگیری ۰/۰۰۰۱ برای بروزرسانی وزن¬های شبکه بکار گرفته شده است. در مرحله آموزش مدل پیشنهادی، اندازه دسته برابر با ۳۲ و اندازه مجموعه اعتبارسنجی برابر با بیست درصد داده‌های آموزشی انتخاب شده است. تعداد مراحل آموزش مدل نیز برابر با ۲۰۰ در نظر گرفته شده است. به منظور جلوگیری از بیش‌برازش، از سازوکار توقف زودهنگام استفاده گردید. بر این اساس، در صورتی که در طول ۱۰ درصد پایانی دورهای آموزشی، بهبودی معناداری در مقدار تابع زیان روی داده‌های اعتبارسنجی مشاهده نشود، فرآیند آموزش به‌طور خودکار متوقف شده و وزن‌های مدل متناظر با کمترین خطای اعتبارسنجی ذخیره می‌گردد.

۴-۴ نتایج

پس از آماده‌سازی داده‌ها و تنظیم پارامترهای مدل‌ها بر اساس توضیحات ارائه‌شده در بخش قبل، مدل‌های رگرسیون سنتی (MLR، KNR و RFR) و مدل‌های MIMO بر روی داده‌های آموزشی برازش داده شدند. سپس، عملکرد هر مدل بر روی داده‌های آزمون ارزیابی گردید. در این بخش، نتایج حاصل از اجرای مدل‌ها ارائه و مقایسه می‌شوند. تحلیل‌ها شامل بررسی دقت پیش‌بینی، ارزیابی شاخص‌های آماری عملکرد و مقایسه کیفی بین مدل‌های مختلف است تا بتوان کارآمدترین روش در پیش‌بینی متغیرهای خروجی کوره گندله‌سازی را شناسایی کرد.

٤-٤-١ نتایج انتخاب ویژگی

در این بخش عملکرد سه الگوریتم انتخاب ویژگی تک متغیره امتیاز اف، اطلاعات متقابل و همبستگی پیرسون روی مجموعه داده کوره گندله‌سازی بررسی خواهد شد. به این منظور، از سه مدل MLR، KNR و RFR استفاده شده است. مدل‌های رگرسیون به ازای هر متغیر خروجی، با ویژگی‌های انتخاب شده توسط هر یک از سه الگوریتم انتخاب ویژگی فیلتری روی نمونه‌های مجموعه داده آموزش، اعمال شده و سپس عملکرد آن‌ها روی مجموعه داده آزمون بررسی خواهد شد. جدول (٤)، میانگین نتایج بدست آمده را برای معیارهای R^2-Score، MAE و MSE را روی ٦٧ متغیر خروجی نشان می‌دهد.

لازم به ذکر است که در این پژوهش، معیارهای ارزیابی عملکرد بر مبنای مقادیر واقعی خروجی‌ها محاسبه شده‌اند. اگرچه دامنه مقادیر متغیرهای دما و فشار متفاوت است، اما هدف از میانگین‌گیری معیارها در اینجا مقایسه نسبی عملکرد مدل‌ها در پیش‌بینی کلیه خروجی‌ها بوده است، نه مقایسه دقیق بین متغیرهای هم‌مقیاس. در واقع، میانگین معیارها تنها برای ایجاد شاخصی تجمیعی جهت مقایسه کیفی بین الگوریتم‌ها به‌کار رفته و تمام مدل‌ها تحت شرایط یکسان آموزش و ارزیابی شده‌اند. بنابراین اختلاف دامنه‌ها بر مقایسه نهایی تأثیرگذار نیست.

بر اساس نتایج گزارش شده در جدول (٤)، از نظر میانگین R^2-Score روی همه متغیرهای خروجی، بهترین دقت برای KNR با الگوریتم انتخاب ویژگی اطلاعات متقابل بدست آمده که برابر با ٨٨/٥٩ درصد است. رتبه‌های بعدی که مقادیر ٨٦/٥٢ درصد و ٨٢/٧٧ درصد هستند، به ترتیب توسط RFR با الگوریتم انتخاب ویژگی اطلاعات متقابل و KNR با امتیاز-اف کسب شده‌اند. همچنین الگوریتم RFR با امتیاز-اف رتبه چهارم R^2-Score را بدست آورده که برابر با ٨١/٩٦ درصد است.

این نتایج نشان می‌دهند که استفاده از الگوریتم انتخاب ویژگی اطلاعات متقابل در مجموع عملکرد بهتری نسبت به دو روش دیگر دارد، به‌ویژه در مدل‌های غیرخطی مانند KNR و RFR. در مقابل، الگوریتم همبستگی پیرسون در اکثر موارد کمترین مقدار R^2-Score را به‌دست آورده است که می‌تواند به محدودیت این روش در شناسایی روابط غیرخطی بین متغیرها نسبت داده شود.

جدول (٤): نتایج میانگین معیارهای ارزیابی مدل‌های رگرسیون روی ٦٧ خروجی در سه روش انتخاب ویژگی

		امتیاز-اف	اطلاعات متقابل	همبستگی پیرسون
MLR	R^2-Score	٥١/٢٥	٤٩/٩	٣٦/٨
	MAE	٨/٧٥	٨/٥٦	٩/٤٥
	MSE	٣٥٥/٨٤	٣٧٥/٥٣	٤١١/٣٤
KNR	R^2-Score	٨٢/٧٧	٨٨/٥٩	٧٨/٣٤
	MAE	٣/٩٥	٣/٠٩	٤/١٨
	MSE	١١٣/٥٨	٨٣/٥٥	١٣١/٥٨
RFR	R^2-Score	٨١/٩٦	٨٦/٥٢	٧٧/٠٢
	MAE	٤/٨١	٤/١	٥/٢
	MSE	١٣٤/٩٣	٩٨/٩٧	١٥٦/٤٤

از نظر معیار MAE، کمترین خطای میانگین مطلق (٣/٠٩) مربوط به مدل KNR با الگوریتم اطلاعات متقابل است، که بیانگر دقت بالاتر پیش‌بینی در این ترکیب است. همچنین در معیار MSE، همین ترکیب با مقدار ٨٣/٥٥ کمترین میانگین خطای مربعات را داشته و در نتیجه بهترین عملکرد کلی را از نظر کاهش خطای پیش‌بینی ارائه داده است. نمودارهای شکل (٥)، مقایسه عملکرد مدل‌های MLR، KNR و RFR را در سه روش انتخاب ویژگی (امتیاز-اف، اطلاعات متقابل، همبستگی پیرسون) از نظر معیارهای R^2-Score، MAE و MSE به صورت بصری نشان می‌دهند.

به طور کلی، می‌توان نتیجه گرفت که در مسئله حاضر، ترکیب انتخاب ویژگی اطلاعات متقابل و مدل KNR بهترین تعادل را بین دقت پیش‌بینی و کاهش خطا ایجاد کرده است. این امر بیانگر آن است که بهره‌گیری از روش‌های انتخاب ویژگی که قادر به شناسایی روابط غیرخطی میان متغیرهای ورودی و خروجی هستند، می‌تواند نقش مؤثری در بهبود عملکرد مدل‌های یادگیری ماشین برای فرآیندهای صنعتی پیچیده ایفا کند.

٤-٤-٢ نتایج مدل‌سازی

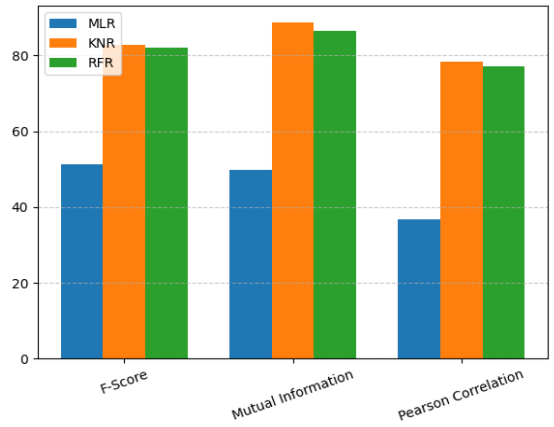
در این بخش، نتایج حاصل از مدل‌سازی با الگوریتم‌های MLR، KNR و RFR بدون مرحله انتخاب ویژگی و با در نظر گرفتن همه متغیرهای ورودی و همچنین مدل‌های MIMO ارائه خواهد شد. جدول (۵) نتایج حاصل از مدل‌سازی را نشان می‌دهد. لازم به ذکر است که در این جدول میانگین معیارهای ارزیابی روی همه خروجی‌ها نشان داده شده است.

جدول (۵): نتایج میانگین معیارهای ارزیابی عملکرد در مدل‌سازی

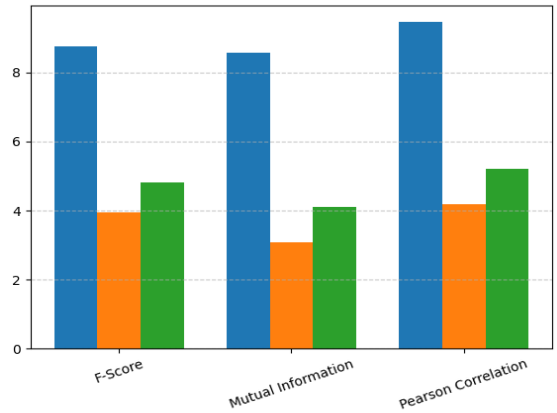
	R2-Score	MAE	MSE
MLR	٦٤/٢٢	٦/٨١	٢٤٣/٨٤
RFR	٨٩/٤١	٣/٥٢	**٥٢/١٢**
KNR	**٩١/٦٧**	**٢/٥**	٧٢/٧٨
MIMO-Temperature	٩١/٥٠	٥/٩٤	١١٤/٩٦
MIMO-Pressure	٨٩/٦٢	٠/٥٤	٠/٨٣
MIMO-All	٨٩/٣٤	٣/٧٦	٧١/٥٦

بر اساس نتایج ارائه‌شده در جدول (۵)، مشاهده می‌شود که عملکرد مدل‌های غیرخطی نسبت به مدل خطی چندمتغیره (MLR) به‌طور محسوسی بهتر است. مدل MLR با مقدار R^2-Score برابر با ٦٤/٢٢ و مقادیر بالای MAE و MSE نتوانسته است ساختارهای پیچیده موجود در داده‌ها را به‌خوبی مدل‌سازی کند و در نتیجه دقت پایینی دارد. در مقابل، دو مدل غیرخطی RFR و KNR توانسته‌اند با افزایش چشم‌گیر ضریب تعیین (به ترتیب ۸۹/٤۱ و ۹۱/٦۷) و کاهش خطاهای مطلق و مربعی، نتایج بسیار بهتری ارائه دهند. در میان این دو مدل، KNR از نظر خطای مطلق عملکرد بهتری دارد، در حالی که RFR در کاهش خطاهای بزرگ و پایدارسازی مدل از طریق MSE، برتری نسبی نشان می‌دهد.

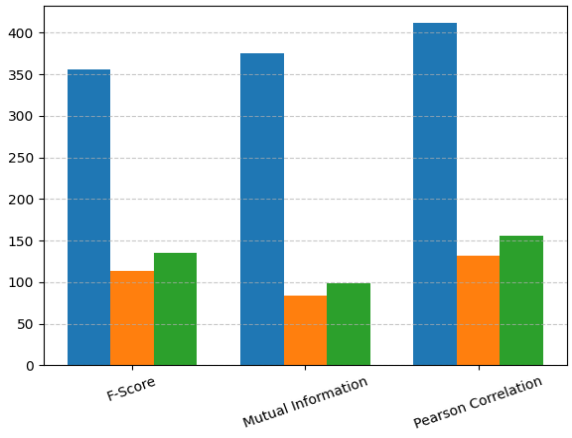
در مقایسه با این مدل‌ها، مدل‌های MIMO نتایج متفاوتی ارائه داده‌اند. مدل MIMO-Temperature از نظر R^2-Score عملکردی مشابه KNR دارد، اما خطاهای آن (MAE و MSE) بالاتر است که می‌تواند ناشی از پیچیدگی پیش‌بینی همزمان چند خروجی باشد. در مقابل، مدل MIMO-Pressure با مقادیر بسیار پایین MAE و MSE (به ترتیب ۰/٥٤ و ۰/۸۳) و R^2-Score بالا، به ظاهر بهترین عملکرد را نشان می‌دهد. با این حال، باید توجه داشت که دامنه تغییرات متغیرهای فشار به‌طور ذاتی بسیار



(الف)



(ب)



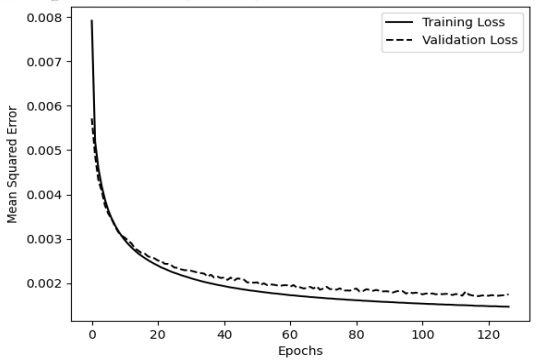
(ج)

شکل (۵): نمودارهای مقایسه عملکرد مدل‌های MLR، KNR و RFR در سه روش انتخاب ویژگی با معیارهای الف) R^2-Score، ب) MAE و ج) MSE

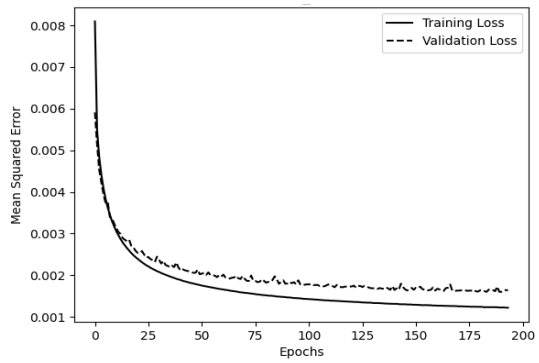
محدودتر از متغیرهای دما است و همین موضوع باعث کوچک شدن مقادیر خطای میانگین و مربعی در مدل MIMO-Pressure شده است. بنابراین، برتری ظاهری این مدل در معیارهای خطا الزاماً به معنای دقت بالاتر آن نسبت به سایر مدل‌ها نیست، بلکه بخشی از این نتایج به تفاوت مقیاس متغیرها باز می‌گردد. به طور کلی، انتخاب مدل باید متناسب با شرایط و محدودیت‌ها باشد:

- اگر منابع محاسباتی محدود باشند و نیاز به پیاده‌سازی سریع وجود داشته باشد، مدل‌های ساده‌ای مانند MLR به دلیل سرعت بالا در آموزش و پیش‌بینی گزینه مناسبی هستند، هرچند دقت پایین‌تری دارند.
- مدل KNR به دلیل نیاز به محاسبات فاصله با تمام داده‌های آموزشی، برای داده‌های بزرگ یا محیط‌های بلادرنگ مناسب نیست، اما در داده‌های کوچک و جایی که دقت بالاتر از سرعت اهمیت دارد می‌تواند کارآمد باشد.
- مدل RFR تعادلی بین دقت و سرعت برقرار می‌کند و نسبت به KNRدر داده‌های بزرگ مقیاس‌پذیرتر است، به همین دلیل در کاربردهای بلادرنگ برتری دارد.
- مدل‌های MIMO مبتنی بر شبکه عصبی به منابع سخت‌افزاری قدرتمند (GPU و RAM بالا) نیاز دارند، اما وقتی چنین منابعی در دسترس باشد و تعداد خروجی‌ها زیاد باشد (مانند دما و فشار کوره)، به دلیل امکان یادگیری روابط پیچیده و پیش‌بینی همزمان چند خروجی گزینه‌ای بهینه محسوب می‌شوند. شکل (٦) نمودارهای تابع زیان مدل‌های MIMO را نشان می‌دهد.

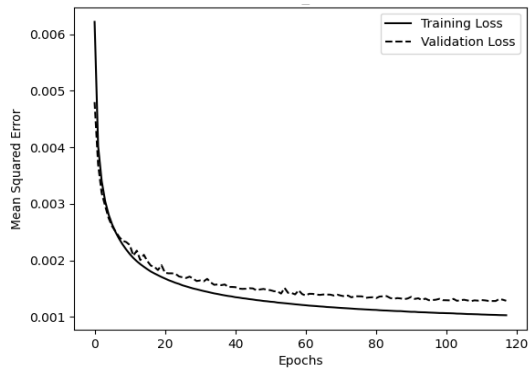
از نظر هزینه محاسباتی، تفاوت قابل‌توجهی میان مدل‌های استفاده‌شده وجود دارد. مدل MLR از نظر پیچیدگی زمانی در حدود $O(N \times n^2 + n^3)$ برای n ویژگی و N نمونه آموزشی است [۳۲]. در عمل رگرسیون خطی معمولاً از نظر زمان اجرا و پیاده‌سازی سریع و سبک است، ولی سرعت آن نسبی است و به مقادیر N و n بستگی دارد. مدل KNR به دلیل نیاز به محاسبه فاصله بین هر نمونه آزمون و تمام داده‌های آموزشی، دارای پیچیدگی زمانی $O(n \times N)$ است [۳۳]. این الگوریتم برای استنتاج بلادرنگ با داده زیاد معمولاً مناسب نیست.



(الف)



(ب)



(ج)

شکل (٦): نمودارهای تابع زیان در سه مدل الف) MIMO_All، ب) MIMO-Temperature و ج) MIMO-Pressure.

در مقابل، RFR با پیچیدگی تقریبی $O(T \times m_{try} \times N \log N)$ است که در آن T تعداد درخت‌ها و m_{try} تعداد ویژگی‌هایی است که در هر گره به صورت تصادفی انتخاب می‌شوند [۳٤]. RFR به‌خاطر قابلیت موازی‌سازی ساخت درخت‌ها در کاربردهای صنعتی یک گزینه متعادل و محبوب است. در مورد مدل MIMO،

بیشترین بار محاسباتی مربوط به آموزش شبکه است که در هر دور آموزشی به‌طور تقریبی $O(E \times N \times P)$ که در آن E تعداد دورهای آموزشی و P تعداد پارامترها هستند. در مجموع، از دیدگاه عملی، مدل‌های RFR و MIMO با توجه به تعادل بین دقت بالا و زمان اجرای قابل‌قبول، گزینه‌های مناسبی برای استفاده در محیط‌های واقعی فرآیند گندله‌سازی محسوب می‌شوند.

۵- نتیجه‌گیری

این پژوهش به مدل‌سازی داده‌محور فرآیند کوره گندله‌سازی مجتمع گل‌گهر با استفاده از الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی پرداخت. روش پیشنهادی شامل سه مرحله اصلی پیش‌پردازش داده‌ها، انتخاب ویژگی و مدل‌سازی بود. نتایج انتخاب ویژگی نشان داد که معیار اطلاعات متقابل به دلیل توانایی در شناسایی روابط غیرخطی، عملکرد بهتری نسبت به امتیاز-اف و ضریب همبستگی پیرسون داشته و دقت مدل‌های غیرخطی مانند RFR و KNR را به‌طور محسوسی ارتقا داده است. در مرحله مدل‌سازی نیز مشخص شد که الگوریتم‌های غیرخطی نسبت به رگرسیون خطی چندمتغیره برتری دارند، به‌گونه‌ای که KNR کمترین خطای مطلق و RFR پایداری بیشتری در کاهش خطاهای بزرگ ارائه کرد. مدل‌های MIMO نیز قابلیت پیش‌بینی همزمان فشار و دما را فراهم کردند. به طور کلی، انتخاب مدل وابسته به شرایط عملیاتی است. مدل‌های ساده مانند MLR برای شرایط با محدودیت منابع محاسباتی مناسب‌اند، RFR برای داده‌های بزرگ و کاربردهای بلادرنگ کارایی بیشتری دارد، و مدل‌های MIMO در صورت دسترسی به منابع سخت‌افزاری قدرتمند بهترین گزینه برای پیش‌بینی همزمان چند خروجی محسوب می‌شوند.

نتایج این تحقیق می‌تواند مبنایی برای توسعه ابزارهای هوشمند کنترل فرآیند در صنعت فولاد باشد. در ادامه، توسعه روش‌های پیشرفته‌تر انتخاب ویژگی برای کاهش عدم قطعیت داده‌ها، بهره‌گیری از شبکه‌های عصبی مدرن نظیر LSTM، GRU و معماری‌های مبتنی بر توجه برای استخراج وابستگی‌های زمانی و مکانی، ترکیب مدل‌های فیزیکی و داده‌محور در قالب رویکردهای ترکیبی، اعتبارسنجی مدل‌ها در محیط واقعی و در نهایت طراحی مدل‌های هوشمندی که علاوه بر پیش‌بینی خروجی‌ها به بهینه‌سازی انرژی و کاهش آلاینده‌ها کمک کنند، از مهم‌ترین مسیرهای آینده این حوزه به شمار می‌روند.

۶- سپاس‌گزاری

این پژوهش برگرفته از پروژه‌ای پژوهشی است که با حمایت شرکت معدنی و صنعتی گل‌گهر انجام شده است. بدین‌وسیله از این شرکت به‌دلیل همکاری، ارائه داده‌های عملیاتی و پشتیبانی فنی در انجام این تحقیق صمیمانه قدردانی می‌شود.

References

[1] F. Habashi, Handbook of extractive metallurgy. Wiley-Vch, 1997.
[2] Saeedi and N. Sotoudeh, Iron Production. Isfahan University of Technology, Jihad Daneshgahi Press, 2006. [In Persian].
[3] M. Alizadeh, M. Alizadeh, and H. Jeylan. “Creating optimal conditions for producing suitable green pellets from iron ore concentrates with low specific surface area using organic and mineral additives,” Journal of Metallurgical Engineering, vol. 21, no. 3, pp. 216–224, 2018. [In Persian].
[4] K. Meier, Y. Nabialahi, and A. Mansouri Aliabadi. Iron Ore Pelletizing. Shamloo Publishing, 2012. [In Persian].
[5] B. Abazarpoor, R. Hejazi, M. Saghaeian, and V. Sheikhzadeh, “Effect of iron ore particle size and shape on green pellet quality,” in International Mineral Processing Congress, Moscow, 2018.
[6] L. Lu, Iron ore: mineralogy, processing and environmental sustainability, Elsevier, 2015.
[7] K. Meyer, Pelletizing of iron ores, Springer-Verlag, 1980.
[8] S. Sadrnezhaad, A. Ferdowsi, and H. Payab, “Mathematical model for a straight grate iron ore pellet induration process of industrial scale,” Computational Materials Science, vol. 44, no. 2, pp. 296-302, 2008.
[9] M. Barati, “Dynamic simulation of pellet induration process in straight-grate system,” International journal of mineral processing, vol. 89, no. 1-4, pp. 30-39, 2008 .
[10] H. Nezamabadipoor, M. Rezaei, and A. Atighi. Statistical and Intelligent Modeling in Engineering and Industry. Shahid Bahonar University of Kerman Press, 2024. [In Persian].
[11] T. Umadevi, P. P. Kumar, P. Kumar, N. F. Lobo, and M. Ranjan, “Investigation of factors affecting

pellet strength in straight grate induration machine," Ironmaking & steelmaking, vol. 35, no. 5, pp. 321-326, 2008.
[12] H. Zhou, Y. He, B. Li, D. Song, Q. Zhu, and Y. Li, "Ironmaking process under artificial intelligence technology: A review," Ironmaking & Steelmaking, 2024.
[13] A. J. Deo, S. K. Behera, and D. P. Das, "Online monitoring of iron ore pellet size distribution using lightweight convolutional neural network," IEEE Transactions on Automation Science and Engineering, vol. 21, no. 2, pp. 1974-1985, 2023.
[14] F. Honpeng and G. Runlong, "Practice of Artificial Intelligence for Iron Ore Pelletizing System," in 2024 5th IEEE Global Conference for Advancement in Technology (GCAT), 2024, pp. 1-4.
[15] Y.-h. Gong, C.-h. Wang, J. Li, M. N. Mahyuddin, and M. T. A. Seman, "Application of deep learning in iron ore sintering process: a review," Journal of Iron and Steel Research International, vol. 31, no. 5, pp. 1033-1049, 2024.
[16] Q. Cao, C. Chi, and J. Shan, "Can artificial intelligence technology reduce carbon emissions? A global perspective," Energy Economics, vol. 143, p. 108285, 2025.
[17] M. M. Carvalho, D. G. Faria, M. G. Pérez, M. Cardoso, and E. K. Vakkilainen, "Review on mathematical models for travelling-grate iron oxide pellet induration furnaces," Energy Procedia, vol. 120, pp. 588-595, 2017.
[18] C. Cavalcanti, G. Wanderley, D. Braga, R. Brito, L. Vasconcelos, and K. D. Brito, "Rigorous modeling of the Traveling Grate stage in the iron ore pellet induration process," Journal of Materials Research and Technology, vol. 31, pp. 3410-3424, 2024.
[19] Y. Wang, Y. Xu, X. Song, Q. Sun, J. Zhang, and Z. Liu, "Novel method for temperature prediction in rotary kiln process through machine learning and CFD," Powder Technology, vol. 439, p. 119649, 2024.
[20] S. Mun and J. Yoo, "Operating Key Factor Analysis of a Rotary Kiln Using a Predictive Model and Shapley Additive Explanations," Electronics, vol. 13, no. 22, p. 4413, 2024.
[21] J. Mourão, M. Huerta, U. de Medeiros, I. Cameron, K. O'Leary, and C. Howey, "Guidelines for selecting pellet plant technology," in Proceedings of the 6th International Congress on the Science and Technology of Ironmaking—ICSTI, 2012.
[22] S. Ding, "Feature selection based F-score and ACO algorithm in support vector machine," in Second International Symposium on Knowledge Acquisition and Modeling, 2009, pp. 19-23.
[23] C.-K. Zhang and H. Hu, "Feature selection using the hybrid of ant colony optimization and mutual information for the forecaster," in International Conference on Machine Learning and Cybernetics, 2005, vol. 3, pp. 1728-1732.
[24] A. Sheikhi. Statistical Analysis and Modeling. Hamrah Elm Publishing, 2018. [In Persian].
[25] L. Breiman, "Random forests," Machine learning, vol. 45, pp. 5-32, 2001.
[26] M. Rezaei and H. Nezamabadi-pour, "A Hybridization Method of Prototype Generation and Prototype Selection for K-NN Rule Based on GSA," Journal of AI and Data Mining, vol. 10, no. 2, pp. 257-268, 2022.
[27] F. Rosenblatt, "The perceptron, a perceiving and recognizing automaton: (Project Para)," Cornell Aeronautical Laboratory, 1957.
[28] M. T. Mitchell, Machine learning, New York: McGraw-hill, 2007.
[29] Y. Bengio, I. Goodfellow, and A. Courville, Deep learning, MIT press Cambridge, MA, USA, 2017.
[30] M. Vazan. Deep Learning: Principles, Concepts, and Approaches. Miad Andisheh Publishing, 2020. [In Persian].
[31] Z. Liu, "A method of SVM with normalization in intrusion detection," Procedia Environmental Sciences, vol. 11, pp. 256-262, 2011.
[32] K. Zhong, P. Jain, and I. S. Dhillon, "Mixed linear regression with multiple components," Advances in neural information processing systems, vol. 29, 2016.
[33] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers-a tutorial," ACM computing surveys (CSUR), vol. 54, no. 6, pp. 1-25, 2021.
[34] G. Louppe, Understanding random forests: From theory to practice. Universite de Liege (Belgium), 2014.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 159-173

Intelligent Process Modeling of a Pelletizing Furnace: A Case Study at Golgohar Mining and Industrial Company

Mohadese Rezaei [1*], Hossein Nezamabadi-pour [2], Reza Khksar-pour[3]

[1] PhD Student, Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran
[2] Professor, Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran
[3] MSc, Golgohar Mining and Industrial Company, Sirjan, Kerman, Iran

Article Information

Original Research Paper

Received:
2025 August 22

Accepted:
2025 October 26

Keywords:
Intelligent Modeling, Machine Learning, Feature Selection, multilayer perceptron neural network, pelletizing furnace

Corresponding Author*:
mo.rezaei@eng.uk.ac.ir

Abstract

The pelletizing process is a key stage in the steel production chain, during which iron ore concentrate is transformed into pellets suitable for use in blast furnaces. The induration (firing) stage within the furnace plays a critical role in determining product quality, as the distribution of temperature and pressure directly affects the physical and mechanical properties of the pellets. The objective of this study is to develop a data-driven model to predict the behavior of the pelletizing furnace at Golgohar Mining and Industrial Company based on real operational data. To this end, process input and output variables were collected and preprocessed, and feature selection was performed using three methods: F-score, mutual information, and Pearson correlation coefficient. Subsequently, Multiple Linear Regression, random forest, k-nearest neighbors, and multilayer perceptron neural network models were trained within a multiple-input–multiple-output (MIMO) framework to simultaneously predict 36 temperature variables and 31 pressure variables. The results indicated that the mutual information-based feature selection method improved the performance of nonlinear models, while the multilayer perceptron achieved the highest accuracy, with an average coefficient of determination of 91.50% for temperature and 89.62% for pressure. These findings demonstrate that data-driven learning approaches can accurately model the complex behavior of the furnace and provide a suitable foundation for designing intelligent control systems and optimizing firing conditions.

: 10.22034/ABMIR.2025.23568.1156

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۸۶-۱۷۵

تحلیل و شناسایی ویژگی‌های کلیدی برای طبقه‌بندی عیوب سطحی صفحات فولادی با استفاده از مدل‌های یادگیری ماشین و روش متعادل‌سازی داده‌ها

حبیب خدادادی [۱*]، سید علی مهری [۲] مهسا خدادادی [۳]

[۱] استادیار، گروه مهندسی برق و کامپیوتر، دانشکده فنی و مهندسی، دانشگاه هرمزگان، بندرعباس، ایران

[۲] مهندس کامپیوتر، گروه کامپیوتر، واحد بندرعباس، دانشگاه آزاد اسلامی، بندرعباس، ایران

[۳] مربی، گروه مهندسی مواد و متالورژی، واحد میناب، دانشگاه آزاد اسلامی، میناب، ایران

چکیده

مقاله پژوهشی

در این پژوهش چالش‌های اصلی در تشخیص عیوب سطحی صفحات فولادی ازجمله نامتوازن‌بودن داده‌ها، هم‌پوشانی ویژگی‌ها و دشواری در شناسایی ویژگی‌های مؤثر مورد بررسی قرارگرفته است. هدف اصلی پژوهش، شناسایی مؤثرترین ویژگی‌ها در افزایش دقت مدل‌های طبقه‌بندی عیوب نظیر خراش، لکه و برجستگی است. برای این منظور سه مرحله تحلیل انجام شد: مرحله اول و دوم شامل اجرای الگوریتم‌های جنگل تصادفی و تقویت گرادیان حداکثری بر داده خام و مرحله سوم شامل اعمال همین الگوریتم‌ها بر داده‌های متوازن‌شده با روش SMOTE بود. نتایج نشان داد که استفاده از SMOTE دقت مدل تقویت گرادیان حداکثری را از ۷۹/۱۸٪ به ۹۱/۷٪ و دقت مدل جنگل تصادفی را از ۸۰/۲۱٪ به ۹۱/۰٪ افزایش داده است. نوآوری اصلی این پژوهش در ترکیب روش‌های یادگیری ماشین با متعادل‌سازی داده و تحلیل عددی اهمیت ویژگی‌هاست. همچنین ویژگی‌های مشترکی چون ویژگی‌های مرتبط با ابعاد هندسی، شاخص‌های مکانی و روشنایی به عنوان شاخص‌های پایدار معرفی شدند که می‌توانند مبنای طراحی سامانه‌های هوشمند کنترل کیفیت در صنعت فولاد باشند. تفاوت قابل توجه در فهرست ویژگی‌های مهم بین الگوریتم‌ها و شرایط داده نشان داد که انتخاب ویژگی در این مسئله امری ثابت نبوده و وابسته به روش یادگیری و نحوه آماده‌سازی داده‌ها است. این رویکرد، علاوه بر ارائه بینش علمی عمیق‌تر در حوزه شناسایی عیوب فولاد، می‌تواند به تصمیم‌گیری دقیق‌تر در انتخاب ویژگی‌های کلیدی در کاربردهای صنعتی مشابه کمک کند.

تاریخ دریافت:
۱۴۰۴/۰۵/۲۹
تاریخ پذیرش:
۱۴۰۴/۰۸/۱۴

کلیدواژه‌ها:
شناسایی ویژگی‌های کلیدی، متعادل‌سازی داده‌ها، عیوب سطحی، صنعت فولاد، یادگیری ماشین

نویسنده‌مسئول:
khodadadi@hormozgan.ac.ir

: 10.22034/ABMIR.2025.23555.1155

E-ISSN: 2821-2037

۱- مقدمه

عیوب سطحی در صفحات فولادی می‌توانند باعث کاهش کیفیت محصول نهایی و در نتیجه کاهش بهره‌وری در فرآیندهای تولید شوند. تشخیص به‌موقع این عیوب نقش مهمی در ارتقاء کنترل کیفیت و بهینه‌سازی هزینه‌ها دارد. با رشد روزافزون داده‌های صنعتی و پیشرفت الگوریتم‌های یادگیری ماشین[1]، امکان تحلیل دقیق‌تر و خودکار عیوب فراهم شده است. این پژوهش تلاش دارد با تمرکز بر شناسایی ویژگی‌های مؤثر از داده‌های مربوط به صفحات فولادی، به طراحی مدلی هوشمند برای طبقه‌بندی نوع عیوب بپردازد.

در سال‌های اخیر، یادگیری ماشین به عنوان یک ابزار قدرتمند در بهبود کیفیت و افزایش بهره‌وری در صنعت فولاد و به‌ویژه در حوزه تشخیص خودکار عیوب سطحی ورق‌های فولادی مطرح شده است. الگوریتم‌های یادگیری ماشین، به‌طور گسترده‌ای برای طبقه‌بندی و تشخیص عیوبی مانند ترک، حباب، اکسیداسیون و ناهمواری سطحی مورداستفاده قرارگرفته‌اند که دارای دقت تشخیص بالایی بوده و زمان بازرسی را به طور چشمگیری کاهش داده‌اند [۱]. ادغام این فناوری‌ها در سیستم‌های بازرسی مبتنی بر بینایی ماشین، نه تنها کاهش خطاهای انسانی را به همراه داشته، بلکه به بهبود کیفیت نهایی محصول و کاهش ضایعات در خط تولید منجر شده است [۲].

تحقیقات و پژوهش در شناسایی خودکار عیوب ورقه‌های فولادی، به طور پیوسته در جریان بوده و هر روز روش‌های موثرتری ارائه می‌شود. در تحقیقات جدید به طور گسترده از روش‌های یادگیری عمیق جهت شناسایی عیوب، استفاده شده است ازجمله در پژوهش سال ۲۰۱۹ [۳]، نویسندگان ضمن معرفی یک مجموعه داده جدید در این حوزه، مدل یادگیری عمیقی به‌نام Reverse Attention، معرفی‌شده است که باعث بهبود تشخیص نقص‌های کوچک می‌شود. پژوهشی دیگر [٤]، به دنبال تشخیص و طبقه‌بندی خودکار نقص‌های سطحی در فولاد نورد گرم است که در این راستا به جای استفاده از مدل‌های شبکه عصبی عمیق CNN از مدل Vision Transformer بهره می‌گیرد و نشان داده می‌شود که استفاده از مدل‌های مبتنی بر ترانسفورمر می‌تواند عملکرد بهتری نسبت به شبکه‌های کانولوشنی سنتی داشته باشد. در مقاله دیگری [۵] در سال ۲۰۲٤، به تشخیص نقص در فولاد با تاکید بر ویژگی‌های لبه[2] پرداخته شده است چرا که نقص‌های فولادی اغلب دارای لبه‌های ضعیف هستند. در این مقاله از روش‌های سنتی استخراج ویژگی لبه مانند Sobel به همراه شبکه‌های عصبی کوچک و سبک برای طبقه‌بندی استفاده شده است. علاوه بر این، استفاده از شبکه عصبی کانولوشنی بهینه‌شده به همراه پیش‌پردازش‌های پردازش تصویر [٦] و استفاده از مکانیزم توجه[3] در معماری شبکه و بهینه‌سازی فرآیند تشخیص چند مقیاسی [۷] نیز جهت تشخیص عیوب سطحی فولاد استفاده شده است. همچنین در پژوهشی دیگری [۸]، چندین روش تشخیص عیوب سطحی فولاد مقایسه شده است و نویسندگان روش‌های مبتنی بر یادگیری عمیق و مجموعه داده‌های استاندارد و نیز معیارهای ارزیابی را مقایسه و تحلیل می‌کنند که می‌تواند به عنوان یک منبع در حوزه کنترل کیفی صنعت فولاد محسوب شود. در رابطه با انتخاب ویژگی‌های برتر و کلیدی در زمینه شناسایی عیوب سطحی فولاد نیز تاکنون کارهای متعددی انجام شده است. در سال ۲۰۲۲ پژوهشی [۹] به ارائه یک روش مبتنی بر یادگیری ترکیبی[4] برای طبقه‌بندی عیوب صفحات فولادی می‌پردازد و چندین روش یادگیری ماشین را در این زمینه مقایسه کرده و درنهایت از ویژگی‌هایی چون ویژگی‌های بافتی و ویژگی‌های شدت نور به عنوان ویژگی‌های برتر و کلیدی در شناسایی عیوب سطحی فولاد نام برده است. همچنین تحقیق دیگری [۱۰] نیز با کمک چندین روش یادگیری ماشین به طبقه‌بندی عیوب فولاد پرداخته و همچنین پیش‌پردازش‌هایی مانند متعادل‌سازی داده‌های آموزشی نیز در نظر گرفته‌شده است و در اینجا نیز ویژگی‌هایی چون مساحت عیب، شکل عیب و شاخص روشنایی عیب به عنوان ویژگی‌های برتر

[1] Machine Learning
[2] Edge Features
[3] Attention Mechanism
[4] Ensemble Learning

شناسایی‌شده است. هر دو پژوهش از داده رایگان Steel Plates Faults[1] [۱۱] استفاده کرده‌اند. دوربنه و همکاران [۱۲]، پوشش جامعی از الگوریتم‌های کلاسیک و دقت آن‌ها را در شناسایی نقص‌های سطحی فولاد ارائه داده‌اند و چندین الگوریتم مورد مقایسه صورت گرفته است؛ اما عدم تحلیل اهمیت ویژگی‌ها و عدم مرتفع سازی عدم تعادل داده‌ها از نقاط ضعف این مقاله است. از دیگر پژوهش‌های انجام شده، مقاله‌ای در سال ۲۰۲۲ توسط فرناندز و همکاران [۱۳] ارائه‌شده است که یک مرور وسیع از کاربرد یادگیری ماشین در تشخیص نقص‌های صنعتی است و در اینجا نشان داده می‌شود که در مسائل عملی نیازمند بررسی‌هایی مثل عدم تعادل داده و مانند آن هستیم.

علی‌رغم پیشرفت‌های قابل توجه در استفاده از روش‌های یادگیری ماشین برای شناسایی عیوب فولاد، چند چالش مهم همچنان باقی است. نخست، توزیع نامتوازن داده‌ها میان کلاس‌های مختلف عیب موجب می‌شود که مدل‌ها به سمت کلاس‌های پرتکرار گرایش پیدا کنند و دقت شناسایی عیوب نادر کاهش یابد. دوم، بسیاری از ویژگی‌های موجود در داده دارای هم‌پوشانی و هم‌بستگی بالا هستند و این امر انتخاب ویژگی‌های مؤثر را دشوار می‌سازد. سوم، مرور پژوهش‌های پیشین نشان می‌دهد که اجماع روشنی در مورد اینکه کدام ویژگی‌ها بیشترین تأثیر را بر دقت مدل دارند وجود ندارد. چهارم، در اغلب مطالعات، تأثیر متوازن‌سازی داده‌ها و انتخاب ویژگی‌ها به صورت عددی و تطبیقی بررسی نشده است. در این پژوهش تلاش شده است تا با ترکیب الگوریتم‌های جنگل تصادفی[2] و تقویت گرادیان حداکثری[3] (XGBoost) و روش متوازن‌سازی داده‌ها، این چالش‌ها به‌صورت نظام‌مند تحلیل و تا حد زیادی برطرف شود.

داده‌های مورداستفاده در این پژوهش از نوع عددی هستند و مقادیر هر ویژگی حاصل پردازش‌های هندسی و نوری از تصاویر خطوط تولید فولاد است. بنابراین هدف پژوهش نه استخراج ویژگی‌های جدید بلکه شناسایی ویژگی‌های کلیدی موجود در داده‌ها برای افزایش دقت مدل است. با توجه به اهمیت موضوع شناسایی ویژگی‌های کلیدی در فرآیند تشخیص عیوب سطحی فولاد و کاستی‌های تحقیقات قبلی، این پژوهش به این موضوع می‌پردازد.

به طور خلاصه نوآوری‌های این پژوهش شامل موارد زیر است:

- شناسایی تفاوت لیست ویژگی‌های مهم در شرایط مختلف: در این پژوهش نشان داده می‌شود که لیست ویژگی‌های مهم وابسته به الگوریتم و روش متعادل‌سازی داده است و ثابت نیست. این موضوع برای حوزه فولاد یک پیام مهم دارد که انتخاب ویژگی‌ها باید متناسب با شرایط داده و مدل انجام شود.
- مقایسه سیستماتیک چند رویکرد انتخاب ویژگی در سه مرحله متفاوت که امکان مقایسه تأثیر متعادل‌سازی داده و الگوریتم انتخاب ویژگی بر نتیجه نهایی را فراهم می‌کند و معرفی مجموعه‌ای از ویژگی‌های پایدار و مشترک میان چند روش مختلف.
- ارائه مقادیر عددی دقیق اهمیت ویژگی‌ها در هر مرحله: در اینجا برعکس بیشتر مقالات که تنها به ارائه نمودار اهمیت ویژگی‌ها بسنده کرده‌اند، برای هر مرحله، جدول عددی اهمیت ویژگی‌ها استخراج‌شده است و مقایسه کمی بین مراحل را ممکن و علمی‌تر می‌کند.
- مقایسه عملکرد مدل‌ها با تمام ویژگی‌ها و با ۱۰ ویژگی برتر و بررسی ترکیب متعادل‌سازی داده‌ها با مدل‌های یادگیری ماشین بر روی داده‌های فولادی.

ادامه مقاله به این صورت سازمان‌دهی شده است: در بخش دوم، الگوریتم‌ها و روش‌های استفاده شده در این مقاله شرح داده می‌شوند. بخش سوم آزمایش‌ها و یافته‌ها مورد بررسی قرار می‌گیرد و مجموعه داده موردنظر جهت آزمایش‌ها معرفی می‌گردد. درنهایت، بخش چهارم به نتیجه‌گیری می‌پردازد.

۲- مواد و روش‌ها

در این مقاله از الگوریتم‌های جنگل تصادفی و تقویت گرادیان حداکثری (XGBoost) جهت طبقه‌بندی عیوب سطحی فولاد

[1] https://archive.ics.uci.edu/dataset/198/steel+plates+faults
[2] Random Forest
[3] Extreme Gradient Boosting (XGBoost)

استفاده شده است و جهت متعادل‌سازی داده‌ها نیز از روش افزایش نمونه‌های مصنوعی اقلیت[1] (SMOTE) استفاده شده است. در ادامه جزئیات این روش‌ها شرح داده‌شده است.

۲-۱ الگوریتم جنگل تصادفی

الگوریتم‌های جنگل تصادفی از قدرتمندترین الگوریتم‌های یادگیری ماشین در حوزه طبقه‌بندی و رگرسیون هستند که از ترکیب تعدادی درخت تصمیم استفاده کرده و ایده «یادگیری تجمیعی» را برای افزایش دقت پیش‌بینی و کاهش بیش‌برازش به کار می‌گیرد.

در این روش، دو مفهوم «نمونه‌برداری بوت استرپ» و انتخاب ویژگی در هر تقسیم نود، مجموعه‌ای از درخت‌های تصمیم متنوع و مستقل تولید می‌شود و نتایج آن‌ها برای تشکیل پیش‌بینی نهایی ادغام می‌شود. نحوه عملکرد این الگوریتم به شرح زیر است [۱۴]:

- ابتدا از مجموعه داده اصلی، چند زیرمجموعه نمونه‌برداری شده با جایگزینی (بوت‌استرپ) ساخته می‌شود.
- هر درخت تصمیم به صورت جداگانه روی یکی از این زیرمجموعه‌ها آموزش داده می‌شود.
- در هر گره تقسیم، تنها بخشی تصادفی از ویژگی‌ها برای پیدا کردن بهترین نقطه تقسیم بررسی می‌شوند، نه همه ویژگی‌ها.
- در انتها، برای مسائل طبقه‌بندی، پیش‌بینی نهایی بر اساس رأی اکثریت درخت‌ها تعیین می‌شود و برای مسائل رگرسیون، میانگین خروجی درخت‌ها گرفته می‌شود.

ازجمله مزایای این الگوریتم، مقاومت بالا در برابر بیش‌برازش[2] نسبت به درخت تصمیم تنها، کارایی مناسب در مسائل با داده‌ها و ویژگی‌های زیاد، قابلیت محاسبه اهمیت هر ویژگی برای تبیین مدل و مناسب برای داده‌های با نویز و داده‌های نامتوازن را می‌توان برشمرد [۱۵].

۲-۲ الگوریتم تقویت گرادیان حداکثری

الگوریتم تقویت گرادیان حداکثری (XGBoost) یکی از روش‌های یادگیری ماشین مدرن است که بر اساس الگوریتم تقویت گرادیان[3] توسعه‌یافته و استفاده بهتری از آن می‌کند و نوعی پیاده‌سازی از درخت تصمیم‌گیری با تقویت شیب است. در اینجا درخت‌های تصمیم به صورت مرحله‌به‌مرحله با هم ترکیب می‌شوند تا یک مدل قوی‌تر ساخته‌شده و در هر مرحله یک مدل جدید با هدف کاهش خطاهای مدل‌های قبلی آموزش داده می‌شود [۱۶].

این روش تأثیر زیادی در بهبود عملکرد مدل‌های یادگیری ماشین داشته و به صورت استاندارد در بسیاری از کاربردها به کار می‌رود. بخشی از کار XGBoost برپایه روش تقویت گرادیان است که ساختار هر مدل جدید بر روی خطاهای مدل‌های قبلی تمرکز می‌کند. این الگوریتم با استفاده از توابع هدف و امکانات بهینه‌سازی کلاس‌های مختلف، مدیریت فرایند آموزش به صورت موازی، مدیریت داده‌های ناقص و استفاده از روش‌های جلوگیری از بیش برازش، قادر به کار با داده‌های کم و کاهش پیچیدگی مدل است. از ویژگی‌های مهم XGBoost می‌توان به موارد زیر اشاره کرد [۱۷]:

- دقت بالا و قابلیت تعمیم‌پذیری خوب به‌واسطه یادگیری مرحله‌به‌مرحله و بهینه سازی دقیق.
- کارایی محاسباتی بالا با استفاده از موازی‌سازی و بهینه‌سازی‌های داخلی.
- مقابله با داده‌های نامتوازن از طریق تنظیم پارامترها.
- امکان ذخیره و بارگذاری سریع مدل‌ها برای کاربردهای عملی.

۲-۳ روش SMOTE

یکی از روش‌های پیشرفته برای مدیریت مشکل عدم تعادل داده‌ها در یادگیری ماشین، به‌ویژه در مسائل طبقه‌بندی، افزایش نمونه‌های مصنوعی اقلیت (SMOTE) است. در بسیاری از مسائل، تعداد نمونه‌ها در کلاس اقلیت بسیار کمتر از کلاس اکثریت است. این عدم تعادل در داده‌ها دقت مدل‌ها را کاهش می‌دهد و باعث تمایل به سمت کلاس اکثریت می‌شود. که این مشکل را با ایجاد نمونه‌های مصنوعی جدید برای کلاس اقلیت کاهش می‌یابد.

[1] Synthetic Minority Oversampling Technique (SMOTE)

[2] Overfitting

[3] Gradient Boosting

اولین معرفی این روش به نوا داتار چلا و همکارانش در سال ۲۰۰۲ نسبت داده می‌شود [۱۸]. این روش عمدتاً به دنبال حل مشکل بیش برازش ناشی از تکرار تصادفی نمونه‌های کلاس اقلیت بود. شیوه کار SMOTE به این صورت است که نمونه‌های جدیدی با استفاده از درون‌یابی خطی بین نمونه‌های کلاس اقلیت و نزدیک‌ترین همسایه‌های آن‌ها در فضای ویژگی ایجاد می‌کند. این امر به افزایش تنوع نمونه‌های کلاس اقلیت کمک کرده و به مدل‌های یادگیری ماشین امکان می‌دهد تا مرزهای دقیق‌تری برای طبقه‌بندی ایجاد کنند. در این روش سه مرحله زیر انجام می‌شود:

۱- برای هر نمونه از کلاس اقلیت، kنزدیک‌ترین همسایه آن در فضای ویژگی‌ها شناسایی می‌شود

۲- نمونه‌های مصنوعی جدید با ترکیب ویژگی‌های نمونه اصلی و یکی از همسایه‌های نزدیک آن به صورت تصادفی بر روی خط بین این دو نمونه ایجاد می‌شوند.

۳- این فرآیند تا جایی ادامه می‌یابد که تعداد نمونه‌های کلاس اقلیت به میزان دلخواه یا متناسب با کلاس اکثریت برسد.

این رویکرد برخلاف روش‌های ساده نمونه‌برداری، باعث افزایش تنوع نمونه‌های اقلیت بدون تکرار مستقیم می‌شود و به بهبود عملکرد مدل‌ها به‌خصوص در داده‌های نامتوازن کمک می‌کند [۱۸].

۳- آزمایش‌ها و یافته‌ها

۳-۱ مجموعه‌داده

در این پژوهش از مجموعه داده Steel Plates Faults [۱۱] که یک داده معتبر در حوزه تشخیص عیوب سطحی فولاد است، بهره گرفته‌شده است. در جدول ۱ نام ویژگی‌ها و معادل فارسی آن را مشاهده می‌کنید. این مجموعه شامل ۱۹۴۱ نمونه از صفحات فولادی است که برای هر نمونه ۲۷ ویژگی عددی و کیفی اندازه‌گیری شده و هدف آن تشخیص نوع عیب سطحی در صفحات فولاد است. داده‌ها در هفت کلاس مختلف برچسب‌گذاری شده‌اند که بیانگر انواع عیوب رایج در فرآیند تولید فولاد هستند.

در این مجموعه داده علاوه بر ویژگی‌های هندسی و نوری پایه نظیر مساحت، محیط و شاخص‌های مکانی، تعدادی ویژگی مشتق‌شده نیز وجود دارد که توسط تهیه‌کنندگان داده طراحی شده‌اند. مانند SigmoidOfAreas و SigmoidOfLuminosity که حاصل اعمال تابع سیگموید بر مقادیر خام هستند تا مقیاس داده‌ها به بازه بین ۰ و ۱ محدود شده و اثر مقادیر پرت کاهش یابد. همچنین، برخی ویژگی‌ها مانند LogOfAreas و Log_X_Index حاصل تبدیل لگاریتمی هستند که با هدف کاهش چولگی توزیع داده و افزایش تفکیک‌پذیری بین نمونه‌ها ایجاد شده‌اند. علاوه بر این، ویژگی‌هایی نظیر Square_Index یا Orientation_Index بیانگر ترکیب‌های هندسی خاصی از متغیرهای اولیه‌اند که اطلاعات ساختاری بیشتری درباره عیوب سطحی فولاد ارائه می‌کنند.

جدول (۱): ویژگی‌های مجموعه‌داده Steel Plates Faults

ردیف	نام ویژگی	معادل فارسی
۱	X_Minimum	حداقل مختصات X
۲	X_Maximum	حداکثر مختصات X
۳	Y_Minimum	حداقل مختصات Y
٤	Y_Maximum	حداکثر مختصات Y
٥	Pixels_Areas	تعداد پیکسل‌های ناحیه
٦	X_Perimeter	محیط در راستای X
۷	Y_Perimeter	محیط در راستای Y
۸	Sum_of_Luminosity	مجموع روشنایی
۹	Minimum_of_Luminosity	حداقل روشنایی
۱۰	Maximum_of_Luminosity	حداکثر روشنایی
۱۱	Length_of_Conveyer	طول نوار نقاله
۱۲	TypeOfSteel_A300	نوع فولاد A300
۱۳	TypeOfSteel_A400	نوع فولاد A400
۱٤	Steel_Plate_Thickness	ضخامت صفحه فولادی
۱٥	Edges_Index	شاخص لبه‌ها
۱٦	Empty_Index	شاخص فضای خالی
۱۷	Square_Index	شاخص مربعی
۱۸	Outside_X_Index	شاخص بیرونی X
۱۹	Edges_X_Index	شاخص لبه در راستای X
۲۰	Edges_Y_Index	شاخص لبه در راستای Y
۲۱	Outside_Global_Index	شاخص کلی بیرونی
۲۲	LogOfAreas	لگاریتم مساحت ناحیه
۲۳	Log_X_Index	لگاریتم شاخص X
۲٤	Orientation_Index	شاخص جهت‌گیری
۲٥	Luminosity_Index	شاخص روشنایی
۲٦	SigmoidOfAreas	سیگموید مساحت‌ها
۲۷	SigmoidOfLuminosity	سیگموید روشنایی

استفاده از این نوع ویژگی‌های مهندسی‌شده در کنار متغیرهای پایه، امکان تحلیل دقیق‌تر و مدل‌سازی مؤثرتر عیوب را فراهم می‌سازد. همان‌طور که بیان شد، در این مجموعه داده، داده‌ها عددی و شامل ویژگی‌های هندسی و نوری استخراج‌شده از تصاویر هستند که در مجموعه داده موردنظر وجود دارد؛ ازاین‌رو نیازی به استخراج ویژگی جدید نبوده و تمرکز مقاله بر انتخاب ویژگی‌های کلیدی بوده است. به همین دلیل، در این پژوهش از الگوریتم‌های یادگیری ماشین کلاسیک استفاده شده است، چرا که این نوع داده‌ها برای مدل‌های شبکه عصبی عمیق مناسب نیستند و یادگیری عمیق عموماً برای داده‌های تصویری خام کاربرد دارد.

در جدول ۲ اسامی نقص‌ها و تعداد رکورد مربوط به هر نقص نشان داده‌شده است. همان‌گونه که از اطلاعات جدول ۲ مشخص است، این مجموعه داده یک مجموعه داده نامتوازن است که ممکن است موجب سوگیری مدل به سمت یادگیری کلاس‌های غالب شود.

جدول (۲): کلاس‌های خروجی (انواع عیوب سطح فولاد)

ردیف	نام عیب	معادل فارسی	تعداد رکورد
۱	Pastry	خمیر مانند (پوسته‌ای)	۱۵۸
۲	Z_Scratch	خط و خش نوع Z	۱۹۰
۳	K_Scatch	خط و خش نوع K	۳۹۱
٤	Stains	لکه‌ها	۷۲
٥	Dirtiness	کثیفی‌ها	٥٥
٦	Bumps	برجستگی‌ها	٤۰۲
۷	Other_Faults	سایر عیوب	٦۷۳

۳-۲ ارزیابی عملکرد مدل‌ها با داده خام

ازآنجاکه بسیاری از مدل‌های یادگیری ماشین از داده‌هایی با برچسب کلاس عددی استفاده می‌کنند، بنابراین قبل از استفاده از این مجموعه داده، ابتدا برچسب‌های رشته‌ای کلاس‌ها را به صورت اعداد ۰ تا ٦ کدگذاری می‌کنیم؛ یعنی نام کلاس هدف مثل Stains, Bumps و غیره به صورت اعداد ۰ تا ٦ تنظیم می‌شود. همچنین در بعضی از نسخه‌های این داده، برچسب کلاس‌ها به صورت ۷ ستون جدا تعبیه شده است به این معنی که در هر ردیف پس از مقادیر ویژگی‌ها، ۷ ستون کلاس اضافه‌شده است که مقادیر همه کلاس‌ها ۰ و فقط کلاس موردنظر ۱ است؛ بنابراین این مورد هم در مجموعه داده، اصلاح‌شده و ۷ ستون به یک ستون با مقادیر ۰ تا ٦ که نماینده یکی از کلاس‌های ذکرشده است، تبدیل می‌شود.

در گام نخست، کلیه تحلیل‌ها و مدل‌سازی‌ها بر روی داده‌های خام انجام شد، بدون آن‌که هیچ‌گونه پیش‌پردازش یا روش‌های افزایش داده به کار گرفته شود. هدف از این مرحله، بررسی توانایی مدل‌های یادگیری ماشین در شناسایی الگوها و عیوب سطحی فولاد بر اساس تمامی ویژگی‌های اولیه بود.

در مرحله ابتدایی پژوهش، به منظور مقایسه عملکرد مدل‌های مختلف یادگیری ماشین، چهار الگوریتم پایه شامل جنگل تصادفی، XGBoost، ماشین بردار پشتیبان[1] و K- نزدیک‌ترین همسایه[2] مورد آزمون بر روی داده خام قرار گرفتند. نتایج اولیه نشان داد که مدل‌های جنگل تصادفی و XGBoost با دقت‌های حدود ۸۰٪، عملکرد بهتری نسبت به مدل‌های SVM و KNN (با دقت کمتر از ٤۰٪) دارند. بر همین اساس، در مراحل بعدی تمرکز بر دو مدل برتر قرار گرفت تا تحلیل دقیق‌تری از اهمیت ویژگی‌ها و نقش متوازن‌سازی داده‌ها ارائه شود.

برای استخراج مهم‌ترین ویژگی‌ها، از الگوریتم‌های جنگل تصادفی و همچنین XGBoost بهره گرفته شد. ازآنجاکه هر کدام از این دو الگوریتم نماینده دو گروه از الگوریتم‌های ترکیبی هستند، از آن‌ها استفاده شد؛ جنگل تصادفی نماینده روش‌های نمونه‌برداری تجمعی[3] و XGBoost نماینده روش‌های تقویت‌کننده[4] هستند. به‌دلیل توانایی هر دو روش در مدل‌سازی داده‌ها و درعین‌حال امکان استخراج اهمیت ویژگی‌ها مورداستفاده قرار گرفتند. هر یک از این الگوریتم‌ها، پس از آموزش بر روی داده‌ها، وزن یا اهمیتی به هر ویژگی اختصاص می‌دهند که بیانگر سهم آن ویژگی در فرآیند تفکیک کلاس‌ها است. به عنوان مثال، درXGBoost این اهمیت‌ها بر اساس کاهش خطای طبقه‌بندی ناشی از استفاده هر ویژگی محاسبه می‌گردد و در جنگل تصادفی این مقادیر با معیارهایی نظیر کاهش ناخالصی جینی[5] یا کاهش خطا در تقسیم‌بندی داده‌ها تعیین می‌شوند. سپس ویژگی‌ها بر اساس مقادیر

[1] Support Vector Machine (SVM)
[2] K-Nearest Neighbors (KNN)
[3] Bagging
[4] Boosting
[5] Gini Impurity Decrease

اهمیت مرتب شدند و ۱۰ ویژگی برتر با بیشترین نقش در بهبود عملکرد مدل انتخاب گردیدند. این فرآیند در حقیقت نوعی انتخاب ویژگی مبتنی بر مدل محسوب می‌شود که ضمن کاهش ابعاد داده، موجب ساده‌تر شدن مدل و کاهش احتمال بیش‌برازش می‌گردد. علاوه بر شناسایی ده ویژگی برتر، مقادیر عددی اهمیت هر ویژگی نیز محاسبه گردید. این مقادیر بیانگر سهم نسبی هر ویژگی در کاهش خطا یا افزایش دقت مدل هستند. به عنوان مثال در مدل جنگل تصادفی ویژگی «LogOfAreas» با اهمیت ۰,۰۷۳ و در مدلXGBoost ویژگی «Log_X_Index» با اهمیت ۰,۳۷۳ به عنوان شاخص‌های کلیدی معرفی شدند. وجود این مقادیر عددی، امکان مقایسه عینی نقش هر ویژگی در مدل را فراهم کرده و نشان می‌دهد کدام متغیرها بیشترین تأثیر را در تمایز بین کلاس‌های عیوب فولادی دارند.

پس از شناسایی ویژگی‌های مهم، دو سناریوی مجزا تعریف شد:

۱- آموزش مدل‌ها با تمامی ۲۷ ویژگی اصلی

۲- آموزش مدل‌ها تنها با ۱۰ ویژگی منتخب

برای هر دو حالت، ابتدا داده‌ها به قسمت‌های آموزش (۸۰ درصد) و آزمایش (۲۰ درصد) تقسیم شد و پس از یادگیری مدل‌ها بر روی داده آموزش، عملکرد مدل‌ها بر اساس دقت طبقه‌بندی[1] (نسبت تعداد نمونه‌هایی که به طور صحیح طبقه‌بندی شده‌اند به تعداد کل نمونه‌ها) بر روی داده آزمایش، مقایسه شد. این مقایسه امکان ارزیابی تأثیر انتخاب ویژگی‌ها بر روی کیفیت پیش‌بینی را فراهم ساخت. در جدول ۳ دقت مدل‌ها و در جدول ٤ و ٥ ویژگی‌های برتر به‌دست آمده در هر مرحله به همراه مقادیر عددی که بیانگر میزان تأثیر هر ویژگی است، نشان داده‌شده است.

جدول (۳): دقت مدل‌ها بر روی داده خام

نام مدل	دقت مدل (درصد)
جنگل تصادفی با تمامی ویژگی‌ها	۸۰/۲۱
جنگل تصادفی با ۱۰ ویژگی برتر	۷۷/۳۸
XGBoost با تمامی ویژگی‌ها	۷۹/۱۸
XGBoost با ۱۰ ویژگی برتر	۷٦/۸٦

جدول (٤): ویژگی‌های برتر به‌دست آمده به همراه مقادیر عددی اهمیت هر ویژگی بر روی داده‌های خام با استفاده از مدل جنگل تصادفی

ردیف	نام ویژگی	مقدار عددی اهمیت ویژگی
۱	LogOfAreas	۰/۰۷۳۲٦۳
۲	Length_of_Conveyer	۰/۰٦۱۱٤٥
۳	Pixels_Areas	۰/۰٥٦٥٦۱
٤	Log_X_Index	۰/۰٥۰۰۱٥
٥	Outside_X_Index	۰/۰٤۷۰٦٦
٦	Steel_Plate_Thickness	۰/۰٤٦۹۱۹
۷	X_Minimum	۰/۰٤٤٤۹٦
۸	Sum_of_Luminosity	۰/۰٤۲۹۰۹
۹	Minimum_of_Luminosity	۰/۰٤۱٦۳٤
۱۰	Edges_Index	۰/۰۳۸۱٥۷

جدول (٥): ویژگی‌های برتر به‌دست آمده به همراه مقادیر عددی اهمیت هر ویژگی بر روی داده‌های خام با استفاده از مدل XGBoost

ردیف	نام ویژگی	مقدار عددی اهمیت ویژگی
۱	Log_X_Index	۰/۳۷۳۲۰٦
۲	TypeOfSteel_A300	۰/۱۳۸٥۹٤
۳	Pixels_Areas	۰/۰٦۹۲۸۲
٤	Steel_Plate_Thickness	۰/۰٥۲۹٥۹
٥	Y_Perimeter	۰/۰٤٦۰٦۸
٦	Orientation_Index	۰/۰٤۰۰٦۰
۷	Length_of_Conveyer	۰/۰۳٤۷٤٤
۸	X_Minimum	۰/۰۲۳۹٥۲
۹	Minimum_of_Luminosity	۰/۰۲۲۲٦۳
۱۰	Square_Index	۰/۰۱۹۹۷۲

نتایج نشان داد که اگرچه دقت مدل‌ها با تمامی ویژگی‌ها اندکی بیشتر بود، اما عملکرد مدل‌ها با ۱۰ ویژگی برتر نیز در سطح مطلوب باقی ماند. این یافته حاکی از آن است که مجموعه‌ای کوچک از ویژگی‌های کلیدی قادر است بخش عمده‌ای از اطلاعات

[1] Accuracy

لازم برای شناسایی عیوب فولادی را فراهم آورد و استفاده از تمامی ویژگی‌ها لزوماً ضروری نیست.

۳-۳ ارزیابی عملکرد مدل‌ها با داده متوازن شده

یکی از چالش‌های اصلی در مجموعه‌داده مورداستفاده، عدم تعادل کلاس‌ها بود؛ به گونه‌ای که برخی از انواع عیوب دارای تعداد نمونه بسیار بیشتری در مقایسه با سایر دسته‌ها بودند. این موضوع می‌تواند منجر به ایجاد سوگیری مدل‌های یادگیری ماشین به سمت کلاس‌های پرتکرار شود و در نتیجه دقت پیش‌بینی برای کلاس‌های کم‌تکرار کاهش یابد. برای رفع این مشکل، در این پژوهش از روش SMOTE استفاده گردید. در این روش، داده‌های مصنوعی برای کلاس‌های اقلیت ایجاد می‌شوند و بدین ترتیب توزیع داده‌ها متوازن می‌شود.

پس از متوازن‌سازی داده‌ها با SMOTE، فرآیند مشابه مرحله اول تکرار گردید. بدین صورت که ابتدا مدل‌های جنگل تصادفی و XGBoost بر روی داده‌های متوازن‌شده آموزش داده شدند. سپس برای شناسایی مهم‌ترین متغیرهای مؤثر در پیش‌بینی نوع عیب، از معیارهای اهمیت ویژگی هر مدل استفاده شد. در فرآیند متوازن‌سازی، داده‌های تولیدشده توسط SMOTE صرفاً در مجموعه آموزش استفاده شده‌اند تا از ایجاد دقت غیرواقعی در ارزیابی جلوگیری شود.

جدول (٦): دقت مدل‌ها بر روی داده متوازن شده

نام مدل	دقت مدل (درصد)
جنگل تصادفی با تمامی ویژگی‌ها	۹۰/۹۹
جنگل تصادفی با ۱۰ ویژگی برتر	۸۹/۵۰
XGBoost با تمامی ویژگی‌ها	۹۱/۷۳
XGBoost با ۱۰ ویژگی برتر	۸۹/٦۱

اهمیت ویژگی‌ها در هر مدل به صورت عددی بیان می‌شود؛ این مقادیر عددی نشان‌دهنده سهم نسبی هر ویژگی در عملکرد مدل هستند و امکان رتبه‌بندی آن‌ها را فراهم می‌کنند. درنهایت، ۱۰ ویژگی برتر با بالاترین اهمیت در هر مدل انتخاب‌شده و نتایج دقت مدل‌ها با همه ویژگی‌ها و با این ۱۰ ویژگی برتر مورد مقایسه قرار گرفت. در جدول ٦ دقت مدل‌ها و در جدول ۷ و ۸ ویژگی‌های برتر به‌دست آمده در هر مرحله به همراه مقادیر عددی که بیانگر میزان تأثیر هر ویژگی است، نشان داده‌شده است.

جدول (۷): ویژگی‌های برتر به‌دست آمده به همراه مقادیر عددی اهمیت هر ویژگی بر روی داده‌های متوازن شده با استفاده از مدل جنگل تصادفی

ردیف	نام ویژگی	مقدار عددی اهمیت ویژگی
۱	Length_of_Conveyer	۰/۰۷۸۱۸۷
۲	LogOfAreas	۰/۰٦۷۸۱۹
۳	Pixels_Areas	۰/۰۵۷۸۷۵
٤	Edges_Index	۰/۰٤۸۷۷۷
۵	Log_X_Index	۰/۰٤۷۳۰۹
٦	Outside_X_Index	۰/۰٤٦۹٤۷
۷	X_Minimum	۰/۰٤۵۵۹۹
۸	Orientation_Index	۰/۰٤۲۸۸۷
۹	Sum_of_Luminosity	۰/۰٤۲۷۲۰
۱۰	Steel_Plate_Thickness	۰/۰٤۱۱٦۳

جدول (۸): ویژگی‌های برتر به‌دست آمده به همراه مقادیر عددی اهمیت هر ویژگی بر روی داده‌های متوازن شده با استفاده از مدل XGBoost

ردیف	نام ویژگی	مقدار عددی اهمیت ویژگی
۱	TypeOfSteel_A400	۰/۲۵۰۳۲۲
۲	LogOfAreas	۰/۲۰۱۲۸۲
۳	Outside_X_Index	۰/۰۹۸۸۲٤
٤	Orientation_Index	۰/۰۵۲۸۱۲
۵	Length_of_Conveyer	۰/۰۵۱۰۷٦
٦	Square_Index	۰/۰۳۵۳۹٦
۷	Steel_Plate_Thickness	۰/۰۳۰۱۱٤
۸	X_Minimum	۰/۰۲۵۲۸۰
۹	Minimum_of_Luminosity	۰/۰۲۳٦۰۸
۱۰	X_Maximum	۰/۰۲۳۱۱۷

برای اطمینان از پایداری مدل‌ها، روش اعتبارسنجی متقابل پنج‌تایی[1] نیز به‌عنوان مبنای ارزیابی تئوریک در نظر گرفته شد و نتایج حاصل میانگین این عملکرد مدل‌ها است.

نتایج این مرحله نشان داد که اعمال روش SMOTE موجب افزایش قابل توجه دقت مدل‌ها گردیده و علاوه بر آن، ویژگی‌های کلیدی‌تر با وضوح بیشتری شناسایی شدند. این امر تأیید می‌کند که متوازن‌سازی داده‌ها پیش‌نیاز مهمی برای تحلیل‌های یادگیری ماشین در داده‌های صنعتی با کلاس‌های نامتوازن است.

[1] 5-Fold Cross Validation

۴-۳ بحث و نتایج

یافته‌های حاصل از مراحل مختلف این پژوهش نشان داد که انتخاب ویژگی و متوازن‌سازی داده‌ها نقش تعیین‌کننده‌ای در بهبود عملکرد مدل‌های یادگیری ماشین برای طبقه‌بندی عیوب صفحات فولادی دارند. در مرحله اول، آموزش مدل‌ها بر روی داده‌های خام صورت گرفت. نتایج نشان داد که با وجود دستیابی به دقت نسبتاً مناسب، همچنان برخی کلاس‌ها به دلیل عدم تعادل داده‌ها با دقت پایینی شناسایی شدند. در این مرحله ویژگی‌هایی همچون LogOfAreas، Length_of_Conveyer، Pixels_Areas، Log_X_Index و Steel_Plate_Thickness به طور مکرر در میان متغیرهای مهم مشاهده شدند که نشان‌دهنده نقش اساسی آن‌ها در شناسایی الگوهای مربوط به عیوب است.

در مرحله دوم، با استفاده از روش SMOTE داده‌ها متوازن شدند. نتایج حاکی از آن بود که دقت مدل‌ها به شکل محسوسی افزایش یافت (بیش از ده درصد) و نشان‌دهنده اهمیت متوازن‌سازی داده‌ها در کاربردهای صنعتی است. علاوه بر بهبود دقت کلی، اهمیت ویژگی‌ها نیز با وضوح بیشتری مشخص گردید. ویژگی‌هایی نظیر LogOfAreas، TypeOfSteel_A400، Length_of_Conveyer، Orientation_Index و Outside_X_Index در این مرحله اهمیت بالاتری یافتند.

تعدادی از ویژگی‌ها تقریباً در تمام نتایج به‌عنوان ده ویژگی برتر ظاهر شدند:

- LogOfAreas (لگاریتم مساحت‌ها)
- Length_of_Conveyer (طول نوار نقاله)
- Log_X_Index (لگاریتم شاخص X)
- Steel_Plate_Thickness (ضخامت صفحه فولادی)
- Luminosity (روشنایی)
- Pixels_Areas (تعداد پیکسل‌های ناحیه)
- Outside_X_Index (شاخص بیرونی X)
- X_Minimum (حداقل مختصات X)

ازآنجاکه این ویژگی‌ها در چندین مدل و شرایط اهمیت بالایی دارند، بنابراین این ویژگی‌ها به احتمال زیاد واقعاً یک عامل بنیادی در ایجاد یا شناسایی عیوب فولادی هستند و می‌توانند به‌عنوان متغیرهای کلیدی برای کنترل کیفیت هوشمند به‌کار گرفته شوند. این تکرار بیانگر پایداری و اهمیت بنیادی این ویژگی‌ها در شناسایی الگوهای عیوب فولادی است، بنابراین به جای استفاده از همه ویژگی‌ها، می‌توان فقط این ویژگی‌های مشترک را مبنای مدل‌سازی قرار داد. این کار باعث کاهش هزینه محاسباتی و افزایش قابلیت تفسیر مدل می‌شود. ویژگی‌های مرتبط با ابعاد هندسی شامل LogOfAreas، Steel_Plate_Thickness، Length_of_Conveyer و X_Minumium نشان می‌دهد که اندازه و ضخامت صفحه فولادی تأثیر مستقیم در احتمال بروز عیب دارد. همچنین ویژگی‌های مرتبط با شاخص‌های مکانی و روشنایی شامل Outside_X_Index، Minimum_of_Luminosity، Log_X_Index، Sum_of_Luminosity و Pixels_Areas بیانگر آن است که مکان قرارگیری عیب و شدت روشنایی تصویر نقش مهمی در شناسایی دقیق عیوب دارند.

یکی از یافته‌های مهم این پژوهش آن است که روش‌های مختلف انتخاب ویژگی، حتی در شرایط مشابه داده‌ای، مجموعه‌های متفاوتی از ویژگی‌های مهم را معرفی می‌کنند. این اختلاف ناشی از تفاوت در ماهیت مدل‌ها و معیارهای ارزیابی اهمیت ویژگی‌هاست؛ به‌طور مثال، جنگل تصادفی بر اساس کاهش خطای جینی یا آنتروپی تصمیم‌گیری می‌کند، در حالی که XGBoost با بهره‌گیری از تقویت گرادیانی، سهم هر ویژگی در بهبود تابع هدف را به‌عنوان معیار اهمیت در نظر می‌گیرد. پیامد این موضوع آن است که انتخاب ویژگی‌های برتر، نسبی بوده و کاملاً وابسته به الگوریتم مورداستفاده است. بنابراین، تکیه بر یک روش منفرد ممکن است باعث از دست رفتن برخی ابعاد مهم داده شود. ازاین‌رو، رویکرد ترکیبی یا مقایسه‌ای میان چند الگوریتم می‌تواند دید جامع‌تری نسبت به عوامل کلیدی مؤثر بر کیفیت سطح فولاد ارائه دهد. این نتیجه همچنین بیانگر آن است که تصمیم‌گیری نهایی در کاربردهای صنعتی (مانند پایش کیفیت و تشخیص عیوب) نباید صرفاً بر اساس خروجی یک مدل صورت گیرد، بلکه نیازمند تحلیل چندجانبه و هم‌پوشانی میان ویژگی‌های منتخب الگوریتم‌های مختلف است.

٤- نتیجه‌گیری و پیشنهادها

در این پژوهش، یک چارچوب سه‌مرحله‌ای برای شناسایی ویژگی‌های کلیدی و بهینه‌سازی مدل‌های طبقه‌بندی در تشخیص عیوب صفحات فولادی ارائه گردید. در گام نخست، با استفاده از الگوریتم جنگل تصادفی و تقویت گرادیان حداکثری بر روی داده‌های خام، ۱۰ ویژگی برتر به ترتیب اهمیت شناسایی و اهمیت عددی هر یک محاسبه شد. نتایج نشان داد که با استفاده از تمامی ویژگی‌ها، دقت مدل‌ها به‌ترتیب ۸۰/۲۱٪ و ۷۹/۱۸٪ بوده و با کاهش ویژگی‌ها به ۱۰ مورد برتر، دقت به ۷۷/۳۸٪ و ۷٦/۸٦٪ کاهش یافت. در گام دوم، اثر متعادل‌سازی داده‌ها با روش SMOTE بر انتخاب ویژگی‌ها و عملکرد مدل‌ها بررسی گردید. در این مرحله، با استفاده از الگوریتم XGBoost + SMOTE دقت مدل با تمام ویژگی‌ها به ۹۱/۷۳٪ و با ۱۰ ویژگی برتر به ۸۹/٦۱٪ رسید. همچنین، الگوریتم Random Forest + SMOTE نیز عملکردی مشابه داشته و دقت ۹۰/۹۹٪ با تمام ویژگی‌ها و ۸۹/۵۰٪ با ۱۰ ویژگی برتر را به دست آورد. مقایسه سه مرحله نشان داد که لیست ویژگی‌های برتر وابسته به نوع الگوریتم و فرآیند پیش‌پردازش داده‌ها است و این مسئله اهمیت تحلیل چندجانبه را در مسائل صنعتی مشابه برجسته می‌کند. همچنین، استفاده از SMOTE باعث بهبود چشمگیر دقت پیش‌بینی در هر دو مدل گردید که بیانگر اهمیت توازن داده در مسائل طبقه‌بندی با توزیع نامتوازن است. در این پژوهش، ویژگی‌های مرتبط با ابعاد هندسی، شاخص‌های مکانی و روشنایی به عنوان ویژگی‌های کلیدی در شناسایی عیوب سطحی صفحات فولاد، انتخاب گردید.

به عنوان پیشنهادی برای کارهای آینده، می‌توان از روش‌های پیشرفته و ترکیبی برای انتخاب ویژگی استفاده کرد و تأثیر نرمال‌سازی و استانداردسازی پیشرفته داده‌ها در شرایط مختلف را بر بهبود عملکرد مدل‌ها در این حوزه بررسی کرد. همچنین پیاده‌سازی چارچوب توسعه‌یافته بر روی داده‌های صنعتی واقعی و بررسی قابلیت تعمیم‌پذیری مدل، از دیگر پیشنهادها است.

References

[1] Y. Zhang, S. Shen, and S. Xu, “Strip steel surface defect detection based on lightweight YOLOv5,” Frontiers in Neurorobotics, vol. 17, Oct. 2023, doi: https://doi.org/10.3389/fnbot.2023.1263739.

[2] Y. Zhan and F. Feng, "Detection and Identification of Strip Surface Defects Based on Deep Learning," 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Guangzhou, China, 2022, pp. 402-405, doi: 10.1109/MLISE57402.2022.00086.

[3] X. Lv, F. Duan, J. Jiang, X. Fu, and L. Gan, “Deep Metallic Surface Defect Detection: The New Benchmark and Detection Network,” Sensors, vol. 20, no. 6, p. 1562, Mar. 2020, doi: https://doi.org/10.3390/s20061562.

[4] V. Vasan, N. V. Sridharan, Sugumaran Vaithiyanathan, and Mohammadreza Aghaei, “Detection and Classification of Surface Defects on Hot-Rolled Steel using Vision Transformers,” Heliyon, vol. 10, no. 19, pp. e38498–e38498, Oct. 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e38498.

[5] Y. Wang et al., “A steel defect detection method based on edge feature extraction via the Sobel operator,” Scientific Reports, vol. 14, no. 1, Nov. 2024, doi: https://doi.org/10.1038/s41598-024-79205-5.

[6] O. C. Han and U. Kutbay, “Detection of Defects on Metal Surfaces Based on Deep Learning,” Applied Sciences, vol. 15, no. 3, p. 1406, Jan. 2025, doi: https://doi.org/10.3390/app15031406.

[7] Y. Jiang, “Surface defect detection of steel based on improved YOLOv5 algorithm,” Mathematical Biosciences and Engineering, vol. 20, no. 11, pp. 19858–19870, Jan. 2023, doi: https://doi.org/10.3934/mbe.2023879.

[8] A. Ashwin, Edmond, B. Gao, and Wai Lok Woo, “A Review and Benchmark on State-of-the-Art Steel Defects Detection,” SN Computer Science, vol. 5, no. 1, Dec. 2023, doi: https://doi.org/10.1007/s42979-023-02436-2.

[9] E. C. ÖZKAT, “A method to classify steel plate faults based on ensemble learning,” Journal of Materials and Mechatronics: A, vol. 3, no. 2, Oct. 2022, doi: https://doi.org/10.55546/jmm.1161542.

[10] A. Kharal, "Explainable artificial intelligence based fault diagnosis and insight harvesting for steel plates manufacturing," arXiv preprint arXiv:2008.04448, 2020.

[11] M. Buscema, S. Terzi, and W. Tastle, "Steel Plates Faults," UCI Machine Learning Repository, 2010. [Online]. Available: https://doi.org/10.24432/C5J88N. [Accessed: Nov. 13, 2025].

[12] A. Dorbane, F. Harrou, and Y. Sun, "Detecting Faulty Steel Plates Using Machine Learning," in Advances in Computing and Data Sciences, S. Silakari, P. P, S. S, and B. Panda, Eds., Cham, Switzerland: Springer Nature, 2024, pp. 321–333. doi: 10.1007/978-3-031-70906-7_27. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-70906-7_27.

[13] M. Fernandes, J. M. Corchado, and G. Marreiros, “Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: a systematic literature review,” Applied Intelligence, vol. 52, no. 12, Mar. 2022, doi: https://doi.org/10.1007/s10489-022-03344-3.

[14] Sadegh Mosharrafzadeh, Bahman Ravaei, and Ehsanollah Koozegar, “Diagnosis of Diabetes Using a Random Forest Algorithm,” JOURNAL OF DIABETES AND METABOLIC DISORDERS, vol. 21, no. 2, pp. 92–100, 2021, [Online]. Available: https://sid.ir/paper/415460/en.

[15] N. Venkatesan and G. Priya, "A Study of Random Forest Algorithm with Implementation Using WEKA," International Journal of Innovative Research in Computer Science and Engineering, vol. 1, no. 6, pp. 156–162, 2015. [Online]. Available:https://www.ioirp.com/Doc/IJIRCSE/i6/JCSE242.pdf.

[16] S. Bakhtiari, " Malware Detection Using Data Mining and Xgboost and Random Forest Algorithms," Journal of Information and communication Technology in policing (JICTP), vol. 3, no. 1, pp. 55-68, Apr. 2020, doi:https://doi.org/10.22034/pitc.2022.1267935.1130.

[17] T. Chen and C. Guestrin, “XGBoost: a Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ’16, vol. 1, no. 1, pp. 785–794, Aug. 2016, doi: https://doi.org/10.1145/2939672.2939785.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” Journal of Artificial Intelligence Research, vol. 16, no. 16, pp. 321–357, Jun. 2002, doi: https://doi.org/10.1613/jair.953.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 175 - 186

Analysis and Identification of Key Features for Classifying Surface Defects in Steel Plates Using Machine Learning Models and Data Balancing Techniques

Habib Khodadadi[1*], Seyed Ali Mehri[2], Mahsa Khodadadi[3]

[1]Assistant Professor, Department of Electrical and Computer Engineering, University of Hormozgan, Bandar Abbas, Hormozgan, Iran

[2] B.Sc. Graduate, Department of Computer, Islamic Azad University, Bandar Abbas, Iran

[3] Lecturer, Department of Materials and Metallurgy, Islamic Azad University, Minab, Iran

Article Information

Original Research Paper

Received:
2025 August 20

Accepted:
2025 November 05

Keywords:
Data Balancing, Key Feature Identification, Machine Learning, Steel Industry, Surface Defects

Corresponding Author*:
khodadadi@hormozgan.ac.ir

Abstract

In this study, the main challenges in detecting surface defects of steel sheets—including data imbalance, feature overlap, and the difficulty of identifying influential features—were investigated. The primary objective of the research was to identify the most effective features for improving the accuracy of defect classification models such as scratch, stain, and bump detection. To achieve this, a three-stage analysis was conducted. The first and second stages involved applying the Random Forest and Extreme Gradient Boosting (XGBoost) algorithms to the raw data, while the third stage applied the same algorithms to data balanced using the SMOTE technique. The results showed that employing SMOTE increased the accuracy of the XGBoost model from 79.18% to 91.7%, and that of the Random Forest model from 80.21% to 91.0%. The main innovation of this study lies in integrating machine learning methods with data balancing and numerical feature-importance analysis. Furthermore, common features such as geometric dimension parameters, spatial indices, and brightness characteristics were identified as stable indicators that can serve as the foundation for designing intelligent quality-control systems in the steel industry. The considerable variation in the list of important features across algorithms and data conditions demonstrated that feature selection in this problem is not fixed but depends on the learning method and data preprocessing approach. This methodology not only provides deeper scientific insights into steel defect detection but also supports more precise decision-making in selecting key features for similar industrial applications.

: 10.22034/ABMIR.2025.23555.1155

E-ISSN: 2821-2037

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال سوم، شماره دوم، پاییز و زمستان ۱۴۰۴، صفحات ۱۹۷-۱۸۷

بهبود بخش‌بندی عیوب سطح فولاد با بهینه‌سازی ابرپارامترهای فیلتر گابور توسط الگوریتم ژنتیک و طبقه‌بندی جنگل تصادفی

پریسا فتحیان بروجنی[۱]، فریبا نمیرانیان[۲*] فهیمه رشیدی[۳]

[۱] دانشجوی ارشد دانشکده فنی، دانشگاه FAU، ارلانگن، آلمان

[۲] مهندس ارشد هوش مصنوعی، شرکت نیکورسام ویرا ایساتیس، ایران

[۳] دانشجوی دکتری دانشکده کامپیوتر، دانشگاه یزد، یزد، ایران

چکیده **مقاله پژوهشی**

کنترل کیفیت سطح محصولات و شناسایی به موقع عیوبی مانند ترک‌ها، حفره‌ها و ناخالصی‌ها در صنعت فولاد از اهمیت بالایی برخوردار است. این عیوب می‌توانند عملکرد مکانیکی، دوام و قابلیت اطمینان فولاد را کاهش دهند. در این پژوهش، رویکردی نوین برای بخش‌بندی و شناسایی هوشمند عیوب سطح فولاد با استفاده از الگوریتم ژنتیک و جنگل تصادفی ارائه شده است. ابتدا ویژگی‌های تصاویر با استفاده از فیلترهای گابور استخراج شده و سپس با الگوریتم جنگل تصادفی، پیکسل‌ها به گروه‌های دارای ویژگی‌های مشابه تقسیم می‌شوند. ابرپارامترهای فیلترهای گابور به‌عنوان ژن‌های کروموزوم در الگوریتم ژنتیک در نظر گرفته شده و با هدف بهینه‌سازی عملکرد بخش‌بندی تنظیم می‌شوند. تابع برازندگی مبتنی بر F1-score دقت طبقه‌بندی را ارزیابی می‌کند. نتایج نشان می‌دهد که بهینه‌سازی ابرپارامترهای فیلترهای گابور توسط الگوریتم ژنتیک، دقت و کارایی بخش‌بندی عیوب سطح فولاد را افزایش داده است. روش پیشنهادی به F1-score برابر ۰/۹۸۰۷ و دقت ۰/۹۸۱۰ دست یافته و نسبت به روش‌های مرجع بهبود عملکرد نشان می‌دهد. این یافته‌ها نشان می‌دهد که روش پیشنهادی می‌تواند به عنوان یک ابزار برای نظارت خودکار بر کیفیت در خطوط تولید فولاد عمل کند.

تاریخ دریافت:
۱۴۰۴/۰۸/۱۹
تاریخ پذیرش:
۱۴۰۴/۰۹/۱۱

کلیدواژه‌ها:
بخش‌بندی سطح فولاد، شناسایی عیوب فولاد، الگوریتم ژنتیک، جنگل تصادفی

نویسنده‌مسئول:
Fariba.namiranian75@gmail.com

doi : 10.22034/ABMIR.2025.23933.1182

E-ISSN: 2821-2037

۱- مقدمه

بخش‌بندی تصویر یکی از فناوری‌های کلیدی در پردازش تصویر و بینایی ماشین است و نقش حیاتی در کنترل کیفیت محصولات فولادی ایفا می‌کند. در این فرآیند، تصاویر سطح فولاد به نواحی مجزا بر اساس ویژگی‌هایی مانند بافت، رنگ و شکل هندسی تقسیم می‌شوند، به‌گونه‌ای که مناطق دارای ویژگی‌های مشابه بتوانند برای شناسایی ترک‌ها، حفره‌ها و حباب‌ها مورد استفاده قرار گیرند [۱]. بخش‌بندی دقیق تصاویر، امکان ارزیابی خودکار و سریع کیفیت فولاد در خطوط تولید را فراهم کرده و از ساخت محصول‌های معیوب جلوگیری می‌کند، که این امر به طور مستقیم بر بهره‌وری و کاهش هزینه‌های تولید تأثیرگذار است.

بخش‌بندی تصویر در صنعت فولاد به طور گسترده‌ای در سیستم‌های بازرسی خودکار خطوط نورد و ریخته‌گری، تشخیص عیوب سطحی در ورق‌ها و شمش‌ها و نظارت بر کیفیت محصول‌های نهایی کاربرد دارد [۴-۲]. انتخاب روش مناسب بخش‌بندی تصاویر سطح فولاد به ویژگی‌های تصویر، شرایط تصویربرداری، میزان نویز و بافت سطح فولاد بستگی دارد [۱]. تاکنون تکنیک‌ها و الگوریتم‌های متعددی برای این منظور ارائه شده‌اند که در ادامه به مرور برخی از آن‌ها پرداخته می‌شود.

در سال ۲۰۱۷ میلادی نئوجی و همکاران روشی کلاسیک برای تشخیص عیوب سطح فولاد ارائه کردند که بر پایه تصویر گرادیان و آستانه‌گذاری تطبیقی استوار است. در این روش، به‌جای استفاده از یک آستانه ثابت برای تفکیک نواحی معیوب، مقدار آستانه به‌صورت پویا با توجه به تعداد پیکسل‌هایی که از سطح خاکستری معینی در تصویر گرادیان فراتر رفته‌اند تنظیم می‌شود. نتایج این مطالعه نشان می‌دهد که این روش از روش‌های ساده آستانه‌گذاری ثابت مثل اوتسو و آستانه‌گذاری محلی بهتر عمل می‌کند [۵].

در سال ۲۰۱۹ میلادی زایلر و همکاران رویکردی مبتنی بر بینایی ماشین برای بازرسی خودکار لبه‌های برشی ورق‌های فولادی ارائه کردند. در این پژوهش یک الگوریتم نوین طراحی شد که تصاویر ثبت‌شده توسط دوربین را به چهار ناحیه‌ی اصلی شامل بخش صیقلی، ناحیه‌ی شکست، ناحیه‌ی شکست معیوب و پلیسه‌ی انتهایی تقسیم‌بندی می‌کند. الگوریتم آن‌ها با ترکیب فیلترگذاری هوموروفیک برای نرمال‌سازی روشنایی، آستانه‌گذاری شدت روشنایی، عملیات مورفولوژیک و روش‌های آماری توانست این بخش‌ها را استخراج کند. این مطالعه با استفاده از ۴۰۰ تصویر صنعتی انجام شد و نشان داد که بخش‌بندی مبتنی بر شدت روشنایی و مرزهای استخراج‌شده می‌تواند مبنایی برای محاسبه شاخص‌های کیفی لبه‌ی برشی و پایش فرآیند در مقیاس صنعتی فراهم کند [۶].

در سال ۲۰۲۲ میلادی اوسامنتیاگا و همکاران به بررسی و مقایسه‌ی روش‌های مختلف تشخیص خودکار عیوب سطح فلزات، به‌ویژه فولاد، پرداختند. در این پژوهش از دو معماری مبتنی بر یادگیری عمیق، شامل YOLOv5 برای آشکارسازی اشیاء و U-Net برای بخش‌بندی معنایی تصاویر استفاده شد. داده‌های آزمایش شامل مجموعه‌داده‌های صنعتی NEU-1800 و NEU-900 بود که دارای عیوب متداول سطح فولاد از جمله ترک، پوسته، حفره، اکسید و فرورفتگی هستند. نتایج آزمایش‌ها نشان داد که مدل YOLOv5 در مقایسه با نسخه‌های قبلی خود دقت مناسبی در تشخیص ناحیه‌ی عیوب داشته و مدل U-Net در تعیین دقیق مرز عیوب عملکرد قابل توجهی از خود نشان داد [۷].

در سال ۲۰۲۳ میلادی ژانگ و همکاران رویکردی مبتنی بر یادگیری عمیق و تطبیق دامنه چندمنبعی برای شناسایی و بخش‌بندی عیوب سطح فولاد ارائه دادند. در این روش از شبکه‌ی پویا سلسله‌مراتبی استفاده شد که با بهره‌گیری از پولینگ پویا و ادغام ویژگی‌های چندمقیاسی، توانست ویژگی‌های بافتی و ساختاری سطوح فولادی را با دقت بیش‌تری استخراج کند. این مدل با به‌کارگیری مکانیزم‌های یادگیری مقابله‌ای[1] و انتقال دانش چنددامنه‌ای[2] قادر بود در داده‌های آزمون عملکرد دقیقی داشته باشد. آن‌ها آزمایش‌های خود را بر روی مجموعه‌داده‌های NEU-Seg و Severstal انجام دادند [۸].

[1] Adversarial learning

[2] Multi-target domain-based knowledge transfer

در سال ۲۰۲٤ میلادی فنگ و همکاران شبکه‌ای به نام FR-Net[1] را برای بخش‌بندی سریع عیوب سطح فولاد نواری معرفی کردند. این مدل بر پایه‌ی ساختار رمزگذار- رمزگشا طراحی شد و با بهره‌گیری از سه ماژول اصلی شامل FFM[2] برای ترکیب ویژگی‌های عمقی و کم‌عمق، FRM[3] برای بهبود بیان ویژگی‌ها، و RASPP[4] برای ادغام اطلاعات چندمقیاسی، دقت و سرعت تشخیص عیوب کوچک را افزایش داد. در آزمایش‌ها روی مجموعه‌داده‌های NEU-Seg و SD-Saliency-900، مدل به‌ترتیب به دقت‌های ۸٤/۵۳ ٪ و ۸۷/۲٤٪ و سرعت ۵۸ فریم بر ثانیه دست یافت که عملکرد بهتری نسبت به مدل‌های متداول نشان داد [۹].

در سال ۲۰۲٤ میلادی راوات و همکاران رویکردی ترکیبی برای طبقه‌بندی عیوب سطح فولاد ارائه کردند. این روش بر پایه **ترکیب** شبکه‌ی عصبی کانولوشنی و جنگل تصادفی طراحی شده است. در این روش، ابتدا شبکه کانولوشنی با چهار لایه کانولوشن و چهار لایه بیشینه‌سازی، ویژگی‌های بافتی سطوح معیوب را در سطوح مختلف استخراج می‌کند. سپس خروجی آن به‌عنوان ورودی به طبقه‌بند جنگل‌تصادفی داده می‌شود تا دقت دسته‌بندی افزایش یابد. آن‌ها مجموعه‌داده‌ای شامل ٤۳٤۸ تصویر از پنج نوع عیب متداول شامل ترک‌ریز، درون‌گیرها، لکه، سطح حفره‌دار و خراش را مورد استفاده قرار دادند. نتایج نشان داد که مدل پیشنهادی با دستیابی به **دقت** ۹۷/۲۸٪ عملکرد پایداری در تشخیص و طبقه‌بندی این عیوب دارد. این پژوهش نشان می‌دهد که ترکیب شبکه کانولوشنی و جنگل تصادفی می‌تواند رویکردی مؤثر برای بهبود دقت سیستم‌های بازرسی خودکار و ارتقای کیفیت تولید در صنایع فولاد باشد [۱۰].

در سال ۲۰۲٤ میلادی گوچلو و همکاران با ارائه مدل deepLabV3+، روشی برای بخش‌بندی عیوب سطح فولاد ارائه کردند. در این پژوهش با بازطراحی ماژول هرمی نمونه‌برداری مکانی گسترش‌یافته و افزودن ماژول فشردگی و تحریک، توانایی مدل در استخراج ویژگی‌های چندمقیاسی و تمرکز بر بخش‌های مهم تصویر افزایش یافت. هم‌چنین با ترکیب ویژگی‌های سطح پایین و سطح بالا، دقت تشخیص مرز عیوب بهبود پیدا کرد. نتایج آزمایش‌ها روی مجموعه‌داده شامل عیوبی مانند خراش، لکه و ناخالصی نشان داد که مدل پیشنهادی نسبت به مدل U-Net عملکرد بهتری داشته و میانگین دقت هم‌پوشانی آن از ۷۶،۸۷ درصد به ۷۹،۰۷ درصد افزایش یافته است [۱۱].

در سال ۲۰۲۵ میلادی وانگ و همکاران روشی مبتنی بر سوپرپیکسل‌ها برای بخش‌بندی عیوب سطح داخلی لوله‌های فولادی ارائه کردند. در این پژوهش، ابتدا تصاویر با درنظرگرفتن اطلاعات لبه به سوپرپیکسل‌های همگن تجزیه می‌شوند. سپس برای توصیف دقیق‌تر بافت نواحی معیوب و سالم از فیلترهای گابور و تبدیل فوریه به‌صورت ترکیبی استفاده شده و ویژگی‌های استخراج‌شده به یک طبقه‌بند ماشین بردار پشتیبان داده می‌شود تا احتمال تعلق هر سوپرپیکسل به ناحیه‌ی معیوب تخمین زده شود. در گام بعد، خروجی ماشین بردار پشتیبان همراه با مدل آمیخته‌گاوسی و وزن‌دهی مبتنی بر لبه‌ها در قالب یک مدل ادغام می‌شود تا مرزبندی دقیق‌تری از عیوب حاصل شود. این روش بر روی ۱۰۰ تصویر صنعتی آزمایش شد و توانست به مقدار IoU برابر ۸۷/۹٤٪ دست یابد [۱۲].

با وجود پیشرفت‌های قابل توجه در روش‌های آستانه‌گذاری کلاسیک، الگوریتم‌های مبتنی بر بینایی ماشین و مدل‌های یادگیری عمیق، مرور ادبیات نشان می‌دهد که هر یک از این رویکردها با محدودیت‌هایی همراه هستند. روش‌های کلاسیک آستانه‌گذاری، هرچند ساده و سریع‌اند؛ اما در مواجهه با بافت‌های غیرهمگن، روشنایی غیریک‌نواخت و مرزهای کم‌کنتراست عملکرد مطلوبی ندارند و به‌سادگی دچار بیش‌تقسیمی یا کم‌تقسیمی می‌شوند. روش‌های مبتنی بر پردازش سنتی تصویر نیز وابستگی زیادی به تنظیم دستی پارامترها و شرایط یکنواخت تصویربرداری دارند. از سوی دیگر، مدل‌های یادگیری عمیق مانند U-Net، FR-Net یا DeepLabV3+ برای دستیابی به عملکرد پایدار نیازمند حجم بزرگی از داده‌های برچسب‌خورده هستند و در کاربردهای صنعتی با کمبود نمونه، تنوع محدود داده یا نویز زیاد دچار بیش‌برازش و

[1] Feature Reused Network (FR-Net)
[2] Feature Fusion Module (FFM)
[3] Feature Refinement Module (FRM)
[4] Residual Atrous Spatial Pyramid Pooling (RASPP)

افت دقت در مرزبندی عیوب می‌شوند. علاوه بر این، بسیاری از مدل‌ها در شناسایی عیوب کوچک، ناپیوسته یا با مرزهای مبهم ضعف دارند. بنابراین نیاز به روشی سبک، کم‌داده‌محور و مقاوم وجود دارد که بتواند مرزبندی دقیق عیوب را انجام دهد. روش این پژوهش با هدف رفع این محدودیت‌ها ارائه شده است.

در این پژوهش با هدف بهبود دقت در شناسایی عیوب سطح فولاد، روشی ترکیبی مبتنی بر الگوریتم ژنتیک و جنگل تصادفی ارائه شده است. در این روش، ویژگی‌های بافتی تصاویر با استفاده از فیلترهای گابور استخراج می‌شوند و سپس به‌وسیله‌ی الگوریتم جنگل تصادفی، پیکسل‌ها بر اساس شباهت ویژگی‌ها گروه‌بندی می‌گردند. به‌منظور دستیابی به بهترین کارایی در بخش‌بندی، پارامترهای فیلتر گابور با استفاده از الگوریتم ژنتیک بهینه‌سازی شده‌اند. نتایج به‌دست‌آمده نشان می‌دهد که این روش موجب افزایش دقت شده است و می‌تواند به‌عنوان رویکردی هوشمند برای کنترل کیفیت خودکار در صنایع فولاد به کار گرفته شود.

ساختار مقاله به این صورت تنظیم شده است: در بخش دوم، مفاهیم پایه‌ای مرتبط با الگوریتم ژنتیک، جنگل تصادفی و فیلتر گابور تشریح می‌شود. در بخش سوم، روش پیشنهادی برای بخش‌بندی و شناسایی عیوب سطح فولاد معرفی می‌گردد. بخش چهارم به تشریح و تحلیل نتایج حاصل از پیاده‌سازی و ارزیابی روش پیشنهادی اختصاص دارد و در نهایت، در بخش پنجم جمع‌بندی و نتیجه‌گیری کلی از پژوهش ارائه می‌شود.

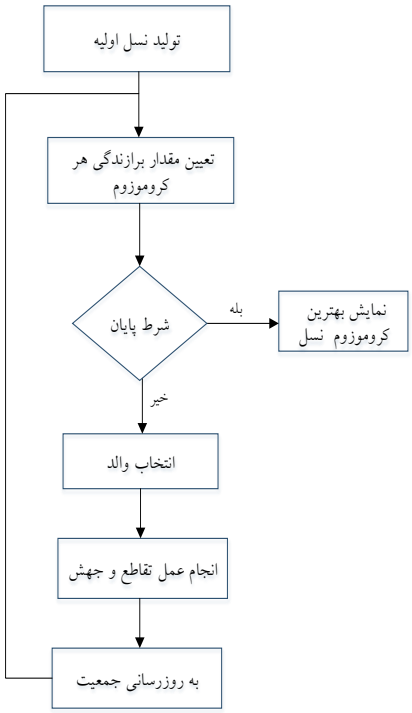
۲- مفاهیم پایه

در این قسمت به معرفی مختصر الگوریتم ژنتیک، الگوریتم جنگل تصادفی و فیلتر گابور که در رویکرد پیشنهادی استفاده شده است، پرداخته می‌شود.

۲-۱- الگوریتم ژنتیک

الگوریتم ژنتیک اولین بار توسط *جان هلند*[1] در سال ۱۹۸۹ مطرح شد و با ایجاد یک مجموعه جواب ابتدایی به عنوان جمعیت اولیه تصادفی آغاز می‌شود. در این الگوریتم پاسخ‌های نامزد به گونه‌ای کدگذاری می‌شوند که بتوان آن‌ها را توسط تابع برازندگی[2] ارزیابی و سپس با عملگرهای ژنتیکی مختلف پردازش کرد. هر نامزد توسط یک کروموزوم[3] نمایش داده می‌شود که شامل رشته‌ای از ژن‌ها[4] است. به جمعیت در هر زمان نسل گفته می‌شود و در هر تکرار الگوریتم، نسل جدید تولید و برازندگی کروموزوم‌ها محاسبه می‌شود.

در تولید نسل بعد، عملگرهای ژنتیک مانند تقاطع[5] و جهش[6] با احتمال مشخص اعمال می‌شوند و نوزادان[7] جدید تولید می‌شوند. برازندگی نوزادان محاسبه شده و کروموزوم‌های بهتر با روش‌های انتخاب به نسل بعد منتقل می‌شوند. اجرای الگوریتم به این صورت است که کروموزوم‌ها با عملگرهای ژنتیک سعی می‌کنند فضای جستجو را بهتر کاوش و برازندگی خود را افزایش دهند تا بهترین راه‌حل‌ها پیدا شوند. برای درک بهتر الگوریتم ژنتیک در شکل ۱ آورده شده است.



شکل (۱): روندنمای الگوریتم ژنتیک

[1] John Holland
[2] Fitness Function
[3] Chromosome
[4] Gene
[5] Cross Over
[6] Mutation
[7] Offspring

در الگوریتم ژنتیک با توجه به کدگذاری مسئله عمل‌گرهای تقاطع و جهش متفاوتی انجام می‌شود. هنگامی که عمل‌گر تقاطع بر روی دو کروموزم اعمال می‌شود، با جابه‌جایی مقادیر به مرور زمان اطلاعات مفید نسل‌های گذشته به کروموزوم‌های جدید منتقل می‌شوند.

در عمل جهش هر ژن با احتمال از پیش تعریف شده، تغییر می‌کند و موجب حذف ژنی از مجموعه ژن‌های جمعیت می‌شود یا ژنی که تا به حال وجود نداشته است به آن اضافه خواهد شد. توسط جهش می‌توان انتظار داشت که کروموزوم‌های خوبی که در مراحل انتخاب و یا تکثیر حذف شده‌اند، دوباره احیا شوند. استفاده از جهش باعث می‌شود عمل کاوش در فضای جواب بهتر صورت گیرد. باید دقت شود فرزندان جدیدی که در هر نسل با استفاده از تقاطع و جهش تولید می‌شوند، بررسی شوند تا محدودیت‌های مسئله را ارضا کنند [۱۳].

۲-۲- الگوریتم جنگل تصادفی

جنگل تصادفی یکی از الگوریتم‌های یادگیری نظارت شده است که برای مسائل رگرسیون[1] و طبقه‌بندی کاربرد دارد. اساس کار این الگوریتم درخت تصمیم است. اگر درخت تصمیم بیش از حد عمیق باشد، ممکن است دچار بیش برازش[2] شود. برای حل این مشکل جنگل تصادفی پیشنهاد شده است [۱۴].

جنگل تصادفی روشی است برای میانگین‌گیری از درخت‌های تصمیم عمیقی که از قسمت‌های مختلف داده آموزشی ایجاد شده باشند. هدف از این روش کاهش واریانس و افزایش عملکرد مدل می‌باشد [۱۵].

۳-۲- فیلتر گابور

فیلترها در پردازش تصویر نقش مهمی دارند و یکی از آن‌ها فیلتر گابور است که به نام دنیس گابور[3] نام‌گذاری شده و یک فیلتر خطی است. این فیلتر برای استخراج ویژگی و تحلیل بافت استفاده می‌شود [۱۶].

رابطه فیلتر گابور دارای دو جزء حقیقی و موهومی است. فیلتر گابور در رابطه ۱ نشان داده شده است.

$$g(x,y;\lambda,\square,\Psi,\sigma,\gamma) = \exp\left(\frac{x^2+\gamma^2y^2}{2\sigma^2}\right)\exp(i\left(2\pi\frac{x}{\lambda}+\Psi\right)) \quad (۱)$$

در این رابطه λ طول موج سینوسی را نشان می‌دهد. $\square$ نشان‌دهنده جهت‌گیری در تابع گابور می‌باشد. Ψ نشان‌دهنده انحراف فاز است. σ انحراف معیار گاوسی و γ نسبت ابعاد و بیضی بودن تابع گابور را مشخص می‌کند.

۳- روش پیشنهادی

هدف اصلی در بخش‌بندی تصاویر سطح فولاد تمایز نواحی مختلف سطح شامل بخش‌های سالم، ترک‌دار، زنگ‌زده و سایر عیوب سطحی است. الگوریتم پیشنهادی این مطالعه در شکل ۲ نمایش داده شده است.

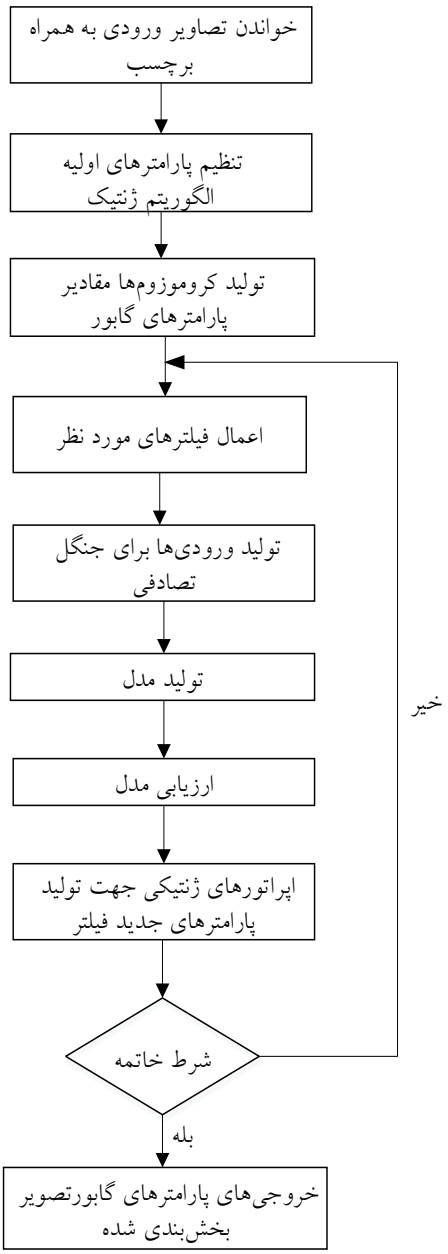
در این روش، تصاویر ورودی ابتدا با مجموعه‌ای از فیلترهای تصویر شامل فیلترهای گابور، کنی[4]، روبرتز[5]، سوبل[6]، اسچر[7]، پریویت[8] و گاوسی کانولوشن می‌شوند تا ویژگی‌های مهم برای بخش‌بندی استخراج گردد. سپس خروجی هر فیلتر به صورت بردار ستونی در نظر گرفته شده و به عنوان ورودی الگوریتم طبقه‌بندی ارائه می‌شود.

با توجه به این که روش مورد استفاده در این تحقیق مبتنی بر یادگیری با نظارت است، تصاویر آموزشی پیش از اجرای الگوریتم توسط کارشناسان خبره بخش‌بندی شده‌اند. این تصاویر مشابه داده‌های ورودی، به بردارهای ستونی تبدیل شده و مراحل پردازش و بخش‌بندی مطابق شکل ۳ قابل مشاهده است.

با توجه به اینکه فیلتر گابور یکی از فیلترهای مؤثر برای استخراج ویژگی‌های مرتبط با بخش‌بندی تصاویر است، در این مطالعه از فیلتر گابور با مقادیر متفاوت ابرپارامترها برای پردازش تصاویر

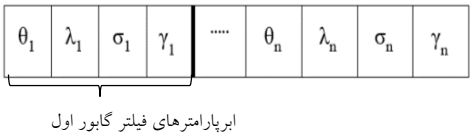
[1] Regression
[2] Overfitting
[3] Dennis Gabor
[4] Canny
[5] Roberts
[6] Sobel
[7] Scharr
[8] Prewitt

استفاده شده است. فیلتر گابور دارای ابرپارامترهای σ ، λ، θو γ است که تنظیم بهینه آن‌ها نقش مهمی در دقت بخش‌بندی دارد.



شکل(۲): روندنمای رویکرد پیشنهادی

در روش پیشنهادی ابرپارامترها با استفاده از الگوریتم ژنتیک به گونه‌ای تعیین می‌شوند که عملکرد بخش‌بندی بهینه گردد. در این الگوریتم، هر کروموزوم نمایانگر مجموعه‌ای از ابرپارامترهای فیلتر گابور است و از آن جا که این پارامترها مقادیر پیوسته دارند، از الگوریتم ژنتیک پیوسته برای بهینه‌سازی استفاده می‌شود. تعداد ژن‌های هر کروموزوم برابر حاصل‌ضرب تعداد فیلترها در تعداد ابرپارامترهای فیلتر گابور است. یک نمونه از کروموزوم در شکل ٤ نشان داده شده است.



شکل (٤): نمونه‌ای از کروموزوم

پس از استخراج ویژگی‌ها، دقت بخش‌بندی تصویر با استفاده از الگوریتم جنگل تصادفی ارزیابی می‌شود و میانگین F1-score به عنوان تابع برازندگی کروموزوم در نظر گرفته می‌شود. فرآیند الگوریتم تا رسیدن به معیار همگرایی از پیش تعیین‌شده ادامه می‌یابد.

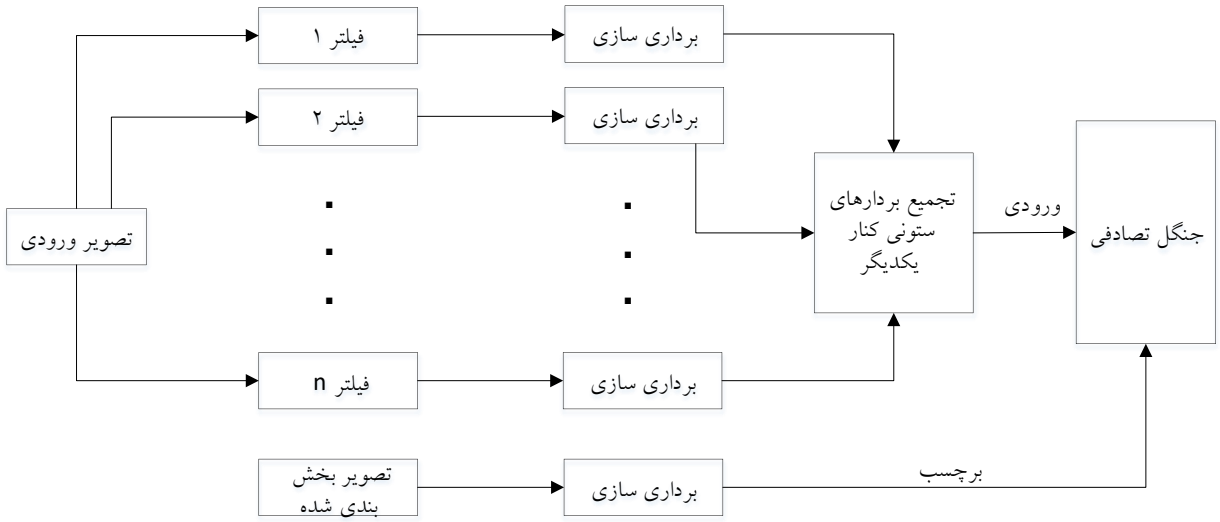
٤- شبیه‌سازی و ارزیابی

در این قسمت به توضیح مجموعه دادگان و مراحل پیاده‌سازی و ارزیابی پرداخته خواهد شد.

٤-١- مجموعه دادگان

در این پژوهش، برای ارزیابی عملکرد الگوریتم بخش‌بندی تصاویر فولادی، از مجموعه‌دادگان Kolektor Surface-Defect (KolektorSDD) استفاده شده است [۱۷]. این مجموعه‌دادگان توسط شرکت Kolektor Group برای شناسایی و تحلیل عیوب سطحی در قطعات فولادی صنعتی تهیه شده و یکی از منابع معتبر در زمینه بازرسی بصری خودکار محسوب می‌شود.

مجموعه KolektorSDD شامل ۳۹۹ تصویر رنگی از سطوح فولادی واقعی است که در شرایط نوری و صنعتی کنترل‌شده ثبت شده است. در این تصاویر، بخشی از نمونه‌ها دارای عیوب سطحی نظیر خراش، ترک‌های ریز، حفره و لکه‌های غیرطبیعی هستند، در حالی که سایر تصاویر مربوط به سطوح سالم و بدون نقص می‌باشند. برای هر تصویر، ماسک دودویی فراهم شده است که ناحیه‌ی دقیق عیب را مشخص می‌کند و به عنوان برچسب برای آموزش و ارزیابی مدل‌های بخش‌بندی مورد استفاده قرار می‌گیرد.



شکل (۳): رویکرد پیشنهادی جهت اعمال فیلتر بر روی تصاویر

استفاده از این مجموعه‌دادگان به دلیل تنوع نوع و شدت عیوب، کیفیت بالای تصاویر و برچسب‌گذاری دقیق، امکان بررسی عملکرد الگوریتم پیشنهادی در تفکیک نواحی معیوب و سالم فولاد را با دقت بالا فراهم می‌سازد. مجموعه دادگان به نسبت ۸۰٪ برای آموزش و ۲۰٪ برای آزمون تقسیم شد. نمونه‌ای از تصاویر این مجموعه به همراه برچسب مربوطه در شکل ۵ نمایش داده شده است.

الف) تصویر اصلی ب) برچسب تصویر

شکل (۵): نمونه‌ای از مجموعه دادگان

۲-٤- پیاده‌سازی و ارزیابی

شبیه‌سازی روش پیشنهادی بر روی یک رایانه با مشخصات حافظه‌ی ۸ گیگابایت، پردازنده‌ی ۷ هسته‌ای و سیستم‌عامل ویندوز ۱۰ انجام شد. پیاده‌سازی الگوریتم‌ها با زبان برنامه‌نویسی پایتون و در محیط توسعه‌ی اسپایدر[1] صورت گرفت.

در شبیه‌سازی روش پیشنهادی از ٤ فیلتر گابور استفاده شد. علاوه بر فیلترهای گابور از فیلترهای کنی، روبرتز، سوبل، اسچر، پریویت، گاوسی و میانه نیز برای استخراج بهتر ویژگی تصاویر استفاده گردید. پارامترهای استفاده شده برای الگوریتم ژنتیک در جدول ۱ آورده شده است.

جدول(۱): پارامترهای الگوریتم ژنتیک در بخش‌بندی تصویر

اندازه جمعیت	۳۰
احتمال تقاطع	۰/۸
احتمال جهش	۰/۰۹
روش انتخاب	چرخ رولت
روش تقاطع	تقاطع یکنواخت
تعداد تکرارها	۵۰

برای عملگر انتخاب در الگوریتم ژنتیک، از چرخ رولت استفاده شده است. در این روش هر کروموزوم معادل با یک قطاع از چرخ

[1] Spyder

دایره‌ای است. اندازه‌ی این قطاع متناسب با برازندگی نرمالیزه هر کروموزوم است. این شیوه‌ی انتخاب بیش‌تر تمایل به انتخاب کروموزوم‌هایی با مقدار برازندگی بالا دارد [۱۸].

مقدار برازندگی مطابق رابطه ۲ از F1-score حاصل از اجرای الگوریتم جنگل تصادفی به دست آمده است.

$$\text{F1-score} = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} \quad (۲)$$

در این رابطه مقدار F1-score از ترکیبی از صحت[1] و پوشش به دست می‌آید. رابطه ۳ و ٤ به ترتیب نشان‌دهنده معیارهای صحت و پوشش است.

$$\text{precision} = \frac{TP}{TP + FP} \quad (۳)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (٤)$$

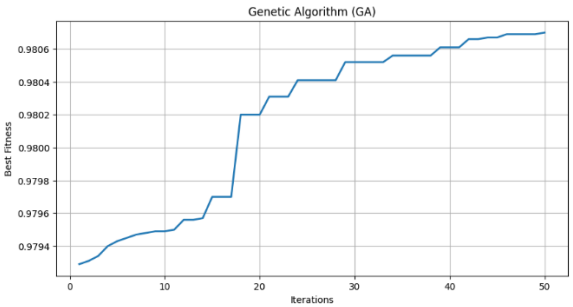
TP نشان‌دهنده پیکسل‌های صحیح بخش‌بندی شده و FP و FN نشان‌دهنده پیکسل‌هایی با بخش‌بندی نادرست هستند. TN نشان‌دهنده پیکسل‌های صحیح بخش‌بندی شده است.

دقت از دیگر معیارهایی است که برای ارزیابی استفاده می‌شود. در رابطه ۵ مقدار دقت آورده شده است.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (۵)$$

در این رابطه دقت برابر است با نسبت بخش‌بندی‌های درست به کل بخش‌بندی‌هایی که انجام شده است.

نمودار همگرایی در شکل ٦ نشان داده شده است. در این نمودار محور افقی تعداد تکرار و محور عمودی بهترین F1-score به دست آمده است.



شکل (۶): نمودار همگرایی الگوریتم ژنتیک بر اساس بهترین مقدار F1-score در طول ۵۰ تکرار

بهترین F1-score به دست آمده برابر با ۰/۹۸۰۷ است که نسبت به مراجع [۱۰] و [۱۲] بهبود قابل توجهی داشته است. مقادیر مناسب به دست آمده برای ابرپارامترهای فیلتر گابور در جدول ۲ آورده شده است.

جدول(۲): مقادیر به دست آمده برای ابرپارامترهای فیلترگابور

فیلتر گابور	σ	θ	λ	γ
فیلترگابور۱	۳/۲۰۷۱٦٤	۱/٦٥۱۲۸	۱/۸۰۲۷٥٤	۰/٦۲۲٤۸۷
فیلترگابور۲	۳/٤٥۳۲۳۲	۰/۰٦۸۱٤	۲/۰۳۳۱٥٤	۰/٤٦۸۹۱٤
فیلترگابور۳	۳/۰۱٤۷٥۱	۱/۷۷۲۲٥	۲/۰٤٦۷۸۱	۰/۳۳٤۱۸۲
فیلترگابور٤	٤/۰۱۳٤۹۱	۲/۱٤٦۹۲	۰/۹٥٤٦۳۱	۰/٤۲٤٥٤۹

به طور خلاصه مقایسه‌ای بین دقت حاصل از روش پیشنهادی و دو روش مراجع [۱۰] و [۱۲] در جدول۳ آمده است.

جدول (۳): مقایسه رویکرد پیشنهادی با رویکردهای پیشین

رویکرد	F1-score	precision	recall	accuracy
رویکرد پیشنهادی	۰/۹۸۰۷	۰/۹۸۲٥	۰/۹۷۹۱	**۰/۹۸۱۰**
رویکرد مرجع [۱۰]	۰/۹۷۱	۰/۹۷٥	۰/۹٦۹	**۰/۹٦٦**
رویکرد مرجع [۱۲]	۰/۹۳٤	۰/۹۱۸	۰/۹٥۲	**۰/۹۷۳**

با توجه به مقایسه انجام گرفته رویکرد پیشنهادی در بخش‌بندی تصویر و شناسایی عیوب فولاد نسبت به دو رویکرد پیشین، نتیجه بهتری را داده است.

۵- نتیجه‌گیری

در این مقاله رویکردی برای بخش‌بندی و شناسایی عیوب سطح فولاد ارائه شد که بر پایه‌ی ترکیب الگوریتم جنگل تصادفی و بهینه‌سازی پارامترهای فیلتر گابور با استفاده از الگوریتم ژنتیک بنا شده است. در روش پیشنهادی، ویژگی‌های بافتی و لبه‌ای تصاویر فولاد با بهره‌گیری از چهار فیلتر گابور و فیلترهای کلاسیکی مانند کنی، روبرتز، سوبل، اسچر، پریویت و گاوسی استخراج گردید. الگوریتم ژنتیک برای تنظیم بهینه‌ی ابرپارامترهای فیلترهای گابور به‌کار گرفته شد، به‌گونه‌ای که هر کروموزوم بیان‌گر مجموعه‌ای از این پارامترها بوده و تابع برازندگی بر اساس میانگین معیار-F1

[1] Precision

score حاصل از عملکرد الگوریتم جنگل تصادفی تعریف گردید. نتایج آزمایش‌ها بر روی تصاویر سطح فولاد نشان داد که روش پیشنهادی در افزایش دقت بخش‌بندی، شناسایی نواحی دارای عیب و تمایز بهتر میان سطوح سالم و معیوب عملکرد مناسبی دارد. این بهبود در F1-score نشان‌دهنده‌ی کارایی بالای روش در بهبود فرآیند کنترل کیفی و کاهش ضایعه‌های تولید در صنایع فولاد است. استفاده از ترکیب هوشمند الگوریتم‌های تکاملی و یادگیری ماشین می‌تواند گامی مؤثر در جهت خودکارسازی بازرسی سطحی فولاد و ارتقای بهره‌وری خطوط تولید باشد.

با وجود عملکرد مناسب روش پیشنهادی، چند مسیر پژوهشی ارزشمند برای توسعه آینده قابل طرح است. نخست آن‌که استفاده از مدل‌های هیبریدی شامل ترکیب استخراج ویژگی‌های کلاسیک با معماری‌های سبک یادگیری عمیق می‌تواند در بهبود دقت و افزایش توانایی مدل در شناسایی عیوب پیچیده مؤثر باشد. دوم، استفاده از روش‌های یادگیری انتقالی و تکنیک‌های افزایش داده می‌تواند چالش کمبود داده‌های برچسب‌خورده صنعتی را تا حدی رفع کند. هم‌چنین توسعه نسخه بلادرنگ الگوریتم جهت کاربرد در خطوط تولید واقعی، ادغام سیستم با سخت‌افزارهای صنعتی و بررسی مقاومت مدل در برابر شرایط متغیر نوری و نویز محیطی از دیگر زمینه‌های مهم پژوهش‌های آتی محسوب می‌شود. در نهایت، پیشنهاد می‌شود عملکرد الگوریتم بر روی مجموعه‌دادگان متنوع‌تر شامل عیوب ریز، ناپیوسته و یا چندکلاسه ارزیابی شود تا قابلیت تعمیم‌پذیری روش در کاربردهای صنعتی گسترده‌تر مورد سنجش قرار گیرد.

References

[1] R. Neven and T. Goedemé, “A multi-branch U-Net for steel surface defect type and severity segmentation,” Metals, vol. 11, no. 6, 2021.

[2] D. Sime, G. Wang, Z. Zeng and B. Peng, “Deep learning-based automated steel surface defect segmentation: a comparative experimental study,” Multimedia Tools Application, vol. 83, no. 1, pp. 2995–3018, 2024.

[3] M. Sharma, L. Jongtae and L. Hansung, “The amalgamation of the object detection and semantic segmentation for steel surface defect detection,” Application Science, vol. 12, no. 12, 2022.

[4] A. Ibrahim and T. Jules, “A survey of vision-based methods for surface defects’ detection and classification in steel products,” Informatics, vol. 11, no. 2, 2024.

[5] N. Neogi, D. Mohanta,P. Dutta, “Defect detection of steel surfaces with global adaptive percentile thresholding of gradient image”, Journal of the Institution of Engineers, vol. 98, no. 6, pp. 557-565,2017.

[6] A. Zeiler, A. Steinboeck, M. Vincze, M. Jochum, ,”Vision-based inspection and segmentation of trimmed steel edges”, IFAC-PapersOnLine, vol. 52, no.14, pp. 165-170, 2019.

[7] R. Usamentiaga, D. Lema, O. Pedrayes and D. Garcia, “Automated surface defect detection in metals: a comparative review of object detection and semantic segmentation using deep learning,” IEEE Transaction Industial Application, vol. 58, no. 3, pp. 4203–4213, 2022.

[8] C. Zhang and X. Zhang, “Multi-target domain-based hierarchical dynamic instance segmentation method for steel defects detection,” Neural Computing Application, vol. 35, no. 10, pp. 7389–7406, 2023.

[9] Q. Feng, F. Li, H. Li and X.Liu, “Feature reused network: a fast segmentation network model for strip steel surfaces defects based on feature reused,” Visual Computer, vol. 40, no. 5, pp. 3633–3648, 2024.

[10] S. Rawat, D. Banerjee, P. Aggarwal and M. Singh, “CNN and random forest fusion for enhanced steel defect classification” , International Conference on Advances in Modern Age Technologies for Health and Engineering Science , pp. 1-6, 2024.

[11] E. Güçlü, İ. Aydın and E. Akın, “Improved deeplabv3+ based steel surface defect segmentation,” in 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems ,pp. 609–613, 2024.

[12] H. Wang, Y. Xie, Y. Wang, C. Zhang, P. Ni, Y. Lu and Y. Wang, “ Surface defect segmentation of steel pipes based on superpixel prior knowledge”, Engineering Research Express, vol. 7, no. 3, 2025.

[13] Z. Shahidi Zandi and A. Latif, “Developing a modern method in circle detection in digital images by using genetic algorithm,” Machine

Vision Image Processing, vol. 8, no. 1, pp. 35–44, 2021.
[14] S. Piryonesi and T. El-Diraby, “Data analytics in asset management: cost-effective prediction of the pavement condition index,” Infrastructre System, vol. 26, no. 1, p. 04019036, 2020.
[15] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo and M. Chica-Rivas, “Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines,” Ore Geology Reviews, vol. 71, pp. 804–818, 2015.
[16] T. Vijayan, M. Sangeetha and A. Kumaravel, “WITHDRAWN: Gabor filter and machine learning based diabetic retinopathy analysis and detection,” Microprocessors and Microsystems, 2020.
[17] “Kolektor Surface-Defect Dataset (KolektorSDD/KSDD).” Accessed: Nov. 04, 2025. [Online]. Available: https://www.vicos.si/resources/kolektorsdd/
[18] A. Lipowski and D. Lipowska, “Roulette-wheel selection via stochastic acceptance,” Physica A: Statistical Mechanics and its Applications, vol. 391, no. 6, pp. 2193–2196, 2012.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 3, No. 2, Autumn & Winter 2025, pp. 187-197

Improving Steel Surface Defect Segmentation by Optimizing Gabor Filter Hyperparameters Using a Genetic Algorithm and Random Forest Classification

Parisa Fathian Broujeni[1], Fariba Namiranian[2*], Fahimeh Rashidy[3]

[1] Master Student, Technische Faculty, FAU University, Erlangen, Germany

[2] AI Engineer, Nikoo Rassam Vira Isatis Company, Iran

[3] PhD student, Faculty of Computer Engineering, Yazd University, Yazd, Iran

Article Information

Original Research Paper

Received:
2025 November 10

Accepted:
2025 December 01

Keywords:
Steel Surface Segmentation, Steel Defect Detection, Genetic Algorithm, Random Forest

Corresponding Author*:
Fariba.namiranian75@gmail.com

Abstract

In the steel industry, it is crucial to control surface quality and to promptly detect defects such as cracks, pores, and impurities, as these flaws can compromise the mechanical performance, durability, and reliability of steel. In this study, a novel approach for image segmentation and intelligent detection of steel surface defects is presented using a combination of the Genetic Algorithm and Random Forest. In the proposed method, image features are extracted using Gabor filters, and then the pixels are classified into groups with similar characteristics through the Random Forest algorithm. The hyperparameters of the Gabor filters are considered as the genes of chromosomes in the Genetic Algorithm and are optimized to enhance the segmentation performance. The fitness function, based on the F1-score, measures the classification performance. The results indicate that optimizing the Gabor filter hyperparameters with the Genetic Algorithm enhances both the accuracy and efficiency of steel surface defect segmentation, achieving an F1-score of 0.9807 and an accuracy of 0.9810, and demonstrating that this method is an effective and reliable tool for automated quality control in steel production lines.

: 10.22034/ABMIR.2025.23933.1182

E-ISSN: 2821-2037

Title **Page**

Smart Steel Industry: A Review of Practical Applications of Artificial Intelligence in Optimization, Sustainability, and Digital Transformation of Production Processes 17

Razieh Sheikhpour

Steel Price Time Series Forecasting Using a Patch-based Transformer Architecture 33

Fatemeh Zare Mehrjardi, Abbas Norouzi

Detection of Steel Bilayer Defects Using Artificial Neural Networks: A Machine Vision and Deep Learning Approach 52

Mostafa Majidi, Amir Makinejad

Smart Energy Optimization and Adaptive Control of Electric Arc Furnaces in the Steel Industry using Digital Twin and Neural Networks 71

Sara Mahmoudi Rashid

Steel Surface Defect Classification Using an Improved VGG16-Based Model 88

Touba Torabipour, Abolfazl Gandomi

A Hybrid Model for Data-Driven Energy Consumption Forecasting in the Steel Industry: An Approach Based on LSTM and Hyperparameter Optimization with the Optuna Algorithm 105

Leila Zolqadr, Majid Shakhsi-Niaei, Yahia Zare Mehrjerdid, Mohammad Mehdi Lotfi

Automatic Detection of Defective Round Steel Billets from Straightening Machine Output Using Machine Vision Techniques 123

Rastin Namiranian, Sahar Soltani, Vali Derhami

SSLA-Net: Semantic–Spatial Learning Architecture for Robust Steel Surface Defect Detection 139

Masoumeh Rezaei, Mansoureh Rezaei, Mehri Rajaei

Intelligent Quality Control and Defect Detection in Steel Manufacturing Processes Using Deep Learning and Image Processing for Energy Optimization 158

Fatemeh Saadatjou, Moslem Molalie

Intelligent Process Modeling of a Pelletizing Furnace: A Case Study at Golgohar Mining and Industrial Company 173

Mohadese Rezaei, Hossein Nezamabadi-pour, Reza Khksar-pour

Analysis and Identification of Key Features for Classifying Surface Defects in Steel Plates Using Machine Learning Models and Data Balancing Techniques 186

Habib Khodadadi, Seyed Ali Mehri, Mahsa Khodadadi

Improving Steel Surface Defect Segmentation by Optimizing Gabor Filter Hyperparameters Using a Genetic Algorithm and Random Forest Classification

Parisa Fathian Broujeni, Fariba Namiranian, Fahimeh Rashidy

In the name of God

Scientific Journal
Applied and basic Machine intelligence research

Print ISSN: 2821-2029
Online ISSN: 2821-2037

Published: Yazd University

Sponsor:
Iranian Alloy Steel Development Company

شرکت توسعه فولاد آلیاژی ایرانیان

ABMIR Office: The Journal Office Address: Department of Computer, Yazd University, Yazd, P.O. Box 89195-741 (Postal Code: 8915818411), IRAN
Phone: 035-31233628
Email: abmir@journals.yazd.ac.ir Website: http://abmir.yazd.ac.ir

Applied and Basic Machine Intelligence Research

Iranian Alloy Steel Development Company

Yazd University

3rd Year – No. 2
Fall and Winter 2025

- **Smart Steel Industry: A Review of Practical Applications of Artificial Intelligence in Optimization, Sustainability, and Digital Transformation...** Razieh Sheikhpour
- **Steel Price Time Series Forecasting Using a Patch-based Transformer Architecture** Fatemeh Zare Mehrjardi, Abbas Norouzi
- **Detection of Steel Bilayer Defects Using Artificial Neural Networks: A Machine Vision and Deep Learning Approach** Mostafa Majidi, Amir Makinejad
- **Smart Energy Optimization and Adaptive Control of Electric Arc Furnaces in the Steel Industry using Digital Twin and Neural Networks** Sara Mahmoudi Rashid
- **Steel Surface Defect Classification Using an Improved VGG16-Based Model** Touba Torabipour, Abolfazl Gandomi
- **A Hybrid Model for Data-Driven Energy Consumption Forecasting in the Steel Industry...** Leila Zolqadr, Majid Shakhsi-Niaei, Yahia Zare Mehrjerdid, Mohammad Mehdi Lotfi
- **Automatic Detection of Defective Round Steel Billets from Straightening Machine Output Using Machine Vision Techniques** Rastin Namiranian, Sahar Soltani, Vali Derhami
- **SSLA-Net: Semantic–Spatial Learning Architecture for Robust Steel Surface Defect Detection** Masoumeh Rezaei, Mansoureh Rezaei, Mehri Rajaei
- **Intelligent Quality Control and Defect Detection in Steel Manufacturing Processes Using Deep Learning and Image Processing for...** Fatemeh Saadatjou, Moslem Molalie
- **Intelligent Process Modeling of a Pelletizing Furnace: A Case Study at Golgohar Mining and Industrial Company** Mohadese Rezaei, Hossein Nezamabadi-pour, Reza Khksar-pour
- **Analysis and Identification of Key Features for Classifying Surface Defects in Steel Plates Using Machine Learning Models...** Habib Khodadadi, Seyed Ali Mehri, Mahsa Khodadadi
- **Improving Steel Surface Defect Segmentation by Optimizing Gabor Filter Hyperparameters Using a Genetic...** Parisa Fathian Broujeni, Fariba Namiranian, Fahimeh Rashidy