

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

عمیق

مصطفی مجیدی^{۱*}، سیدامیر مکی‌نژادسانج^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 ارائه شده است. عیب دوپوستگی که به دلیل 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 به صورت 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 داده شده است. این سامانه شامل سه دوربین صنعتی با زاویه ۱۲۰ درجه و نورپردازی کنترل شده است که امکان تصویربرداری پیوسته و همگام با حرکت خط تولید را فراهم می‌کند. در مرحله اول از تکنیک‌های پیش پردازش تصویر نظیر نرمال سازی و قطعه بندی برای بهبود کیفیت تصویر استفاده شده است. سپس مدل‌های یادگیری عمیق، برای طبقه بندی و شناسایی عیب دوپوستگی آموزش داده شده‌اند. 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 به عنوان جایگزینی مؤثر برای روش‌های سنتی بازرسی عمل کند.

تاریخ دریافت:

۱۴۰۴/۰۵/۳۱

تاریخ پذیرش:

۱۴۰۴/۰۸/۲۲

کلیدواژه‌ها:

تشخیص عیب فولاد، دوپوستگی،

شبکه‌های عصبی مصنوعی،

یادگیری عمیق، بینایی ماشین

نویسنده مسئول:

m.majidi@stu.yazd.ac.ir

doi : 10.22034/ABMIR.2025.23565.1159

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، گزارش شده است که تقاضای جهانی فولاد در سال‌های اخیر با افزایش چشم‌گیری روبه‌رو 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 عدم تشخیص 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^۱ 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 بیل‌ها 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 پایداری 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^۲ 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

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^۱ و ماتریس هم‌وقوع سطوح خاکستری^۲ اشاره کرد. ویژگی‌های حاصل از این روش‌ها معمولاً به‌عنوان ورودی به الگوریتم‌های یادگیری ماشین مانند SVM یا KNN مورد استفاده قرار می‌گیرند [۶-۸]. برخی از روش‌های دیگر بافت از تحلیل سیگنال تصویر در حوزه فرکانس با فیلترهایی مانند Gabor و Wavelet استفاده می‌کنند.

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^۱ 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 ۴ 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 ۶ 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 سستی 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 [۴]. 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.

³ Gray-Level Co-occurrence Matrix (GLCM)

¹ End-to-End

² Local Binary Pattern (LBP)

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 روش‌ها به دنبال این بودند که شبکه‌ای طراحی کنند که با تأکید بر مدل‌های از پیش آموزش دیده و استخراج ویژگی‌های 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^۴ 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 نه تنها 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 برجسته‌خورده در محیط‌های واقعی امکان‌پذیر نیست؛ زیرا تولید چنین داده‌هایی در خطوط تولید صنعتی هزینه‌بر و زمان‌گیر است. این چالش موجب رشد روش‌های بدون نظارت شد. در این رویکرد، مدل‌ها با استفاده از تصاویر بدون عیب آموزش

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 تجزیه و تحلیل 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 بهینه‌سازی 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^۱ 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^۲ که به توصیف لبه یا مرز بیرونی ناحیه شیء می‌پردازند و دوم، توصیف‌گرهای ناحیه‌ای^۳ 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

³ Region-based descriptors

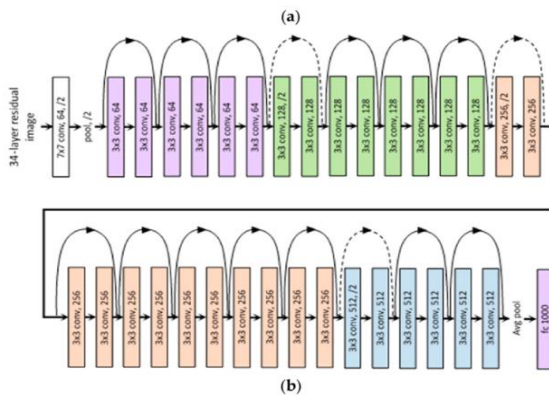
⁴ Ensemble Learning

¹ Shape Descriptors

² Contour-based descriptors

۳-۱ ResNet

معماری ResNet با معرفی اتصالات باقیمانده^۵ در شبکه‌های عصبی عمیق، تحولی بنیادین در حوزه یادگیری عمیق ایجاد کرد. این اتصالات به شبکه اجازه می‌دهند تا اطلاعات و گرادین‌ها را از لایه‌های ابتدایی مستقیماً به لایه‌های عمیق‌تر منتقل کند و بدین ترتیب، مشکل ناپدید شدن گرادین^۶ را که یکی از چالش‌های اصلی در آموزش شبکه‌های بسیار عمیق است، برطرف سازد. در نتیجه، معماری ResNet امکان 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 داده شده است.



شکل (۱): معماری مدل ResNet [۱۶]

۳-۲ EfficientNet

مدل EfficientNetB0 به عنوان نسخه پایه از خانواده EfficientNet در سال ۲۰۱۹ توسط تیم گوگل معرفی شد. این معماری بر پایه مفهوم مقیاس‌گذاری ترکیبی طراحی شده است که در آن سه عامل اصلی شبکه عمیق، عرض و رزولوشن ورودی به صورت هم‌زمان و متعادل افزایش می‌یابند. این رویکرد سبب می‌شود مدل با حداقل تعداد پارامتر، به حداکثر دقت ممکن در طبقه‌بندی تصاویر دست یابد. در نسخه B0، شبکه از بلوک‌های MBConv^۷ استفاده می‌کند که پیش‌تر در معماری MobileNetV2 معرفی شده بودند و به واسطه ساختار سبک و

می‌بینند تا بتوانند هرگونه انحراف از الگوی نرمال را به عنوان 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 بدون نظارت به تنهایی توان پاسخ‌گویی کامل به نیازهای پیچیده و پویای محیط‌های صنعتی را نداشتند. از این‌رو، نسل جدیدی از روش‌ها تحت عنوان یادگیری نیمه‌نظارتی^۱ و یادگیری ضعیف‌نظارتی^۲ 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^۳ و شبکه‌های مولد^۴ ارائه کردند که قادر بود ویژگی‌های ظریف و غیرخطی عیوب سطحی را به صورت خودکار استخراج کند. همچنین، ژانگ و همکاران در مدل CADN [۱۵]، تنها با استفاده از برچسب‌های سطح تصویر توانستند به طور 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 به سوی یادگیری داده‌محور و کم‌نمونه هموار کرد.

اغلب روش‌های موجود بروی داده‌ها و دسته‌بندی‌های شناخته شده 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 استفاده شده است که در ادامه مختصری از آن‌ها بیان می‌شود.

^۵ Residual Connections

^۶ Vanishing Gradient Problem

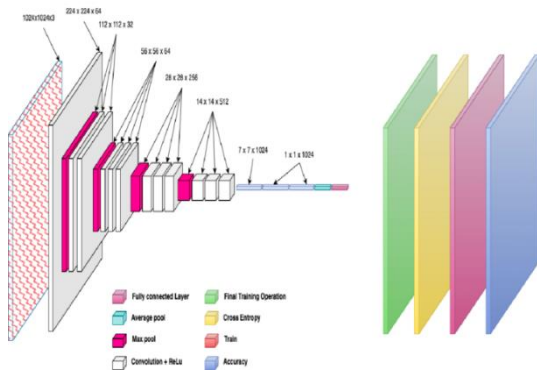
^۷ Mobile Inverted Bottleneck Convolution

^۱ Semi-Supervised Learning

^۲ Weakly-Supervised Learning

^۳ Autoencoders

^۴ Generative Networks



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 ماشینی 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 طراحی شده‌اند. 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

کارآمد خود، باعث کاهش قابل‌توجه حجم محاسبات می‌شوند. افزون بر این، این مدل با بهره‌گیری از بهینه‌سازی خودکار معماری^۱ 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^۲ و دستگاه‌های دارای منابع محاسباتی محدود طراحی شده است. نوآوری اصلی این معماری در به‌کارگیری پیچش‌های جداپذیر عمقی^۳ نهفته است؛ رویکردی که عملیات پیچش متداول را به دو مرحله مجزا تقسیم می‌کند، به گونه‌ای که در مرحله نخست پیچش عمقی^۴ هر فیلتر را به‌صورت مستقل بر روی یک کانال ورودی اعمال می‌کند و در مرحله دوم پیچش نقطه‌ای^۵ 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 داده شده است. در این پژوهش از MobileNetV2 که یک نسخه بهبودیافته از MobileNet است استفاده می‌کنیم.

۳-۴ 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 استفاده شده 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.

^۴ Depthwise Convolution

^۵ Pointwise Convolution

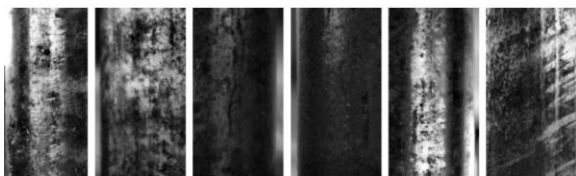
^۱ Neural Architecture Search

^۲ Embedded Systems

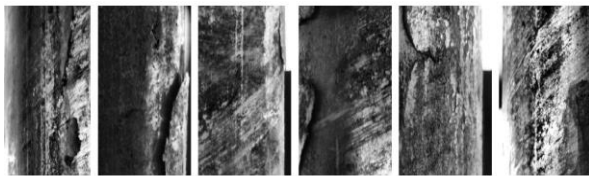
^۳ Depthwise Separable Convolutions

- تصاویر سیاه و سفید: تمامی تصاویر به صورت تک کاناله ثبت شده‌اند. 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 تضمین شده 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 قرار گرفته و به صورت دستی برچسب‌گذاری می‌شوند. در این فرآیند، هر تصویر در یکی از دو کلاس دپوسته (معیوب) یا غیردپوسته (سالم) قرار می‌گیرد. این مرحله برای ایجاد مجموعه داده‌ای 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 ۶ نمایش داده شده‌اند.



شکل (۵): نمونه تصویر میلگرد فولادی سالم.



شکل (۶): نمونه تصویر میلگرد فولادی با عیب دپوسته.

این مجموعه داده به دو بخش آموزش و آزمون تقسیم شده 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 گرفته شده است.

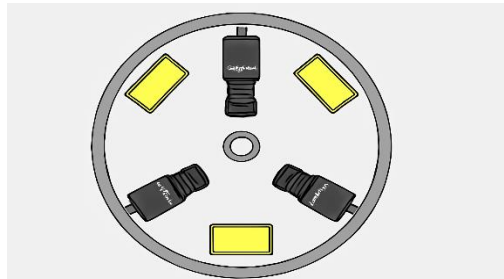
گزینه‌ای ایده‌آل برای کاربردهای بینایی ماشین در محیط‌های صنعتی تبدیل می‌کند.



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 استوانه‌ای استفاده شده 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 ثبت شده وجود عیب دپوستگی 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 استفاده شده است تا یک محیط ایستا و با نور کنترل شده ایجاد شود. این تنظیمات موجب کاهش نویز نوری و بهبود کیفیت ویژگی‌های استخراج شده توسط مدل می‌گردد. آرایش کلی اجزای سیستم تصویربرداری به صورت شماتیک در شکل (۴) نشان داده شده است.

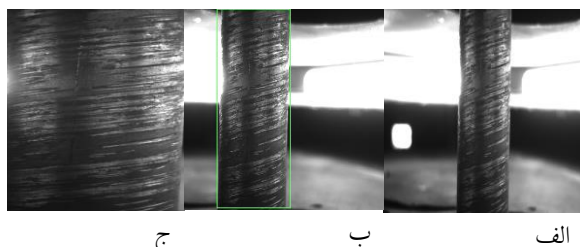


شکل (۴): نحوه چینش دوربین و چراغ‌ها در محیط

بر همین اساس، ویژگی‌های مجموعه تصاویر گردآوری شده به صورت زیر خلاصه می‌شود:

روشنایی را در تصویر محاسبه کرده و خطوط و لبه‌های میلگرد را به‌وضوح نمایان می‌سازد.

در صورتی که در خروجی الگوریتم سوپل هیچ لبه‌ای شناسایی نشود، نتیجه گرفته می‌شود که میلگرد هنوز وارد ناحیه دید دوربین نشده است. در این حالت، تصویر مربوط 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 شناسایی شده تولید می‌گردد. این ماسک بر روی تصویر اصلی اعمال شده 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 برش خورده به عنوان ورودی به مراحل بعدی پردازش ارسال خواهد شد.



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

۴-۲-۲ تغییر اندازه تصویر^۳

پس از جداسازی میلگرد، تصاویر به اندازه یکنواخت ۲۲۴×۲۲۴ پیکسل تغییر اندازه داده می‌شوند. این اندازه استاندارد ورودی برای بسیاری از مدل‌های پیش‌آموزش‌دیده شبکه‌های عصبی پیچشی (مانند VGG, ResNet, Inception) است [۶]. تغییر اندازه یکنواخت، تضمین می‌کند که تمامی تصاویر ورودی به مدل دارای ابعاد یکسان باشند.

۴-۲-۳ نرمال‌سازی^۴

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

جدول (۱): توزیع داده‌های آموزشی و آزمون

بخش داده	کلاس سالم	کلاس دوپسته (معیوب)	جمع کل
آموزش	۱۹۳۴	۱۶۳۶	۳۵۷۰
آزمون	۲۰۰	۲۰۰	۴۰۰

همان‌گونه که مشاهده می‌شود، مجموعه داده 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.

۴-۲-۴ 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:

۴-۲-۱ تشخیص لبه یا قطعه‌بندی میلگرد^۱

در ابتدا، بخش مربوط به میلگرد باید از پس‌زمینه جدا شود. این کار می‌تواند از طریق روش‌های تشخیص لبه (مانند فیلتر سوپل، کنی یا لاپلاس) [۴] 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^۲ 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 سوپل استفاده می‌شود. این الگوریتم تغییرات شدت

^۳Resizing

^۴Normalization

^۱Segmentation

^۲Background Image

این پژوهش، از شبکه‌های عصبی پیچشی^۷ استفاده شده 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 قرار گرفته‌اند. معماری‌های منتخب شامل سه شبکه پیش‌آموزش‌دیده یعنی MobileNetV2، ResNet50 و EfficientNetB0 و یک شبکه Custom CNN طراحی شده 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.

۱-۵ مدل پیش‌آموزش‌دیده

مدل‌های MobileNetV2، ResNet50 و EfficientNetB0 که پیش‌تر بر روی مجموعه داده بزرگ ImageNet آموزش‌دیده‌اند، در این پژوهش به‌عنوان استخراج‌کننده‌های ویژگی مورد استفاده قرار گرفتند. به‌منظور سازگاری این مدل‌ها با مسئله حاضر، دو لایه انتهایی هر شبکه آزادشده^۸ تا در فرآیند تنظیم دقیق^۹، آموزش یابند، درحالی‌که سایر لایه‌ها ثابت^{۱۰} باقی مانده‌اند تا ویژگی‌های عمومی حاصل از مرحله پیش‌آموزش حفظ گردند.

پس از بخش استخراج ویژگی، یک لایه تمام‌متصل با ۶۴ نورون و تابع فعال‌سازی ReLU اضافه 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.

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 پایداری مدل در طول آموزش کمک می‌کند و از تأثیر مقیاس‌های مختلف ویژگی‌ها بر فرآیند یادگیری جلوگیری می‌کند.

۴-۲-۴ داده‌افزایی^۱

برای افزایش حجم و تنوع مجموعه داده آموزشی و جلوگیری از بیش‌برازش^۲، 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 برجستگی آن‌ها تغییر کند. تبدیلات مورد استفاده در این پژوهش عبارت‌اند از:

- چرخش^۳: چرخش تصاویر در زوایای مختلف.
- قرینه‌سازی^۴: قرینه‌سازی افقی یا عمودی تصاویر.
- بزرگ‌نمایی^۵: بزرگ‌نمایی یا کوچک‌نمایی بخش‌هایی از تصویر.
- تغییر روشنایی و کنتراست^۶: ایجاد تغییرات جزئی در روشنایی و کنتراست.
- برش تصادفی^۷: حذف تصادفی بخش‌های غیرضروری تصویر.
- افزودن نویز^۸: 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

⁷Random Cropping

⁸Adding Noise

⁹Convolutional Neural Networks(CNNs)

¹⁰Unfrozen

¹¹Fine-Tuning

¹²Frozen

¹Data Augmentation

²Overfitting

³Rotation

⁴Flipping

⁵Zooming

⁶Brightness and Contrast Adjustment

۲-۵ مدل پیش‌بینی ساده

در کنار مدل‌های پیش‌آموزش‌دیده، یک شبکه Custom CNN نیز به صورت 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 پیش‌بینی متوالی است که هر بلوک از یک لایه کانولوشن دوبعدی و یک لایه ادغام حداکثری تشکیل شده است. تعداد فیلترها در لایه‌های پیش‌بینی به ترتیب ۳۲، ۶۴ و ۱۲۸ در نظر گرفته شده است تا با افزایش عمق شبکه، قابلیت استخراج ویژگی‌های پیچیده‌تر از داده‌های تصویری فراهم شود.

خروجی لایه‌های پیش‌بینی پس از تبدیل به بردار مسطح به یک لایه تمام‌متصل با ۱۲۸ نورون و تابع فعال‌سازی ReLU متصل می‌شود. به منظور کاهش خطر بیش‌برازش، یک لایه Dropout با نرخ ۰.۶ به کار گرفته شد. در نهایت، مشابه سایر مدل‌ها، یک لایه خروجی سیگموئید برای طبقه‌بندی دودویی اضافه گردید. در تحلیل‌ها، این مدل با نماد Custom_CNN نمایش داده شده است.

۳-۵ آموزش و ارزیابی مدل

آموزش مدل‌های یادگیری عمیق مستلزم استفاده از مجموعه داده‌ای 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 مجموعه داده، داده‌افزایی و تنظیمات آموزش، معیارهای ارزیابی مدل‌ها تشریح می‌شود.

۱-۳-۵ تنظیمات آموزش

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

جدول (۲): تنظیمات آموزش مدل‌های یادگیری عمیق.

پارامتر	مقدار
ابعاد ورودی تصویر	۲۲۴×۲۲۴ پیکسل
اندازه دسته	۳۲
تعداد دوره‌های آموزشی	۵۰
نرخ یادگیری	۰/۰۰۰۰۰۵
نرخ Dropout	۰/۶
توقف زودهنگام	پس از ۱۰ گام بدون بهبود
بهینه‌ساز	Adam
تابع زیان	آنترپی متقاطع دودویی ^۱

تابع زیان آنترپی متقاطع دودویی برای طبقه‌بندی دودویی به کار گرفته شده است، زیرا اختلاف بین احتمال پیش‌بینی شده توسط مدل و برچسب واقعی را به صورت مستقیم اندازه‌گیری می‌کند. بهینه‌ساز Adam نیز به دلیل سرعت همگرایی بالا و پایداری در آموزش شبکه‌های عمیق انتخاب شده است. همچنین از زمان‌بندی نرخ یادگیری تطبیقی استفاده شده تا در مراحل پایانی آموزش، نرخ یادگیری به تدریج کاهش یابد و مدل به حداقل بهینه همگرا شود. به منظور پیاده‌سازی سیستم جاری از زبان برنامه‌نویسی پایتون و چارچوب Tensorflow استفاده شده است. همچنین برای توسعه و اجرای کدها از Google Colab با ۱۵ گیگابایت GPU و ۱۶ گیگابایت رم استفاده شده 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.

۲-۳-۵ 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 استفاده شده است. این معیارها شامل موارد زیر هستند:

- **دقت^۲:** نسبت کل پیش‌بینی‌های صحیح به تعداد کل نمونه‌ها. در معادله ۱ فرمول آن نشان داده شده است.

²Accuracy

¹Binary Cross-Entropy

فرآیند آموزش و ارزیابی مدل‌ها برای تحلیل بهتر رفتار آن‌ها در طول آموزش بررسی شده‌اند.

۱-۶ نتایج کمی

در جدول ۳ مقادیر نهایی معیارهای ارزیابی برای مدل‌های مختلف آورده شده است. مطابق نتایج مشاهده شده، مدل MobileNetV2 بالاترین عملکرد را در میان مدل‌های بررسی شده 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 قرار گرفته است، در حالی که 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	۰/۷۳۷۴	۰/۸۱۲۰
EfficientNetB0	۰/۵	۰/۲۵۰۰	۰/۵۰۰۰	۰/۳۳۳۳	۰/۲۹۱۶
MobileNetV2	۰/۹۱۲۵	۰/۹۱۶۳	۰/۹۱۲۵	۰/۹۱۲۳	۰/۹۷۳۱
Custom_CNN	۰/۹	۰/۹۰۲۶	۰/۹۰۰۰	۰/۸۹۹۸	۰/۹۶۷۵

در مقایسه عملکرد مدل‌ها، نتایج نشان داد که دو مدل ResNet50 و EfficientNetB0 دقت پایین‌تری نسبت به سایر مدل‌ها دارند. در مدل EfficientNetB0 عملکرد ضعیف (دقت ۰/۵۰ و مقدار AUC برابر 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 طراحی شده، برای دستیابی به تعمیم مناسب نیازمند داده‌های متنوع‌تر و تنظیم دقیق پارامترهای یادگیری است. به همین دلیل، مدل دچار کم‌پرازش شده و در تفکیک صحیح نمونه‌های سالم و معیوب ناکام مانده است. در مقابل، مدل ResNet50 با 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

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

- **صحت^۱:** نسبت مثبت‌های واقعی به کل پیش‌بینی‌های مثبت. در معادله ۲ فرمول آن نشان داده شده است.

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

- **یادآوری^۲:** نسبت نمونه‌های مثبت واقعی که به درستی توسط مدل شناسایی شده‌اند به کل نمونه‌های مثبت واقعی موجود در داده. در معادله ۳ فرمول آن نشان داده شده 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 ۴ فرمول آن نشان داده شده است.

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (۴)$$

- **منحنی ROC^۳ و مساحت زیر منحنی^۴:** 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].

۶- نمودارها و نتایج

در این بخش، نتایج حاصل از آموزش مدل‌های مختلف شبکه عصبی پیچشی و ارزیابی عملکرد آن‌ها بر اساس معیارهای استاندارد از جمله دقت، صحت، یادآوری، امتیاز F1 و مساحت زیر منحنی (AUC) ارائه می‌شود. علاوه بر این، نمودارهای مربوط به

^۳Receiver Operating Characteristic

^۴Area Under Curve(AUC)

^۱Precision

^۲Recall

دچار خطاهای متعدد شده و نتوانسته ویژگی‌های بافتی مرتبط با عیب دوپوستگی را به‌درستی یاد بگیرد. این مسئله عمدتاً ناشی از کم‌برازش به دلیل 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، به گونه‌ای که توزیع پیش‌بینی‌ها در ماتریس درهم‌ریختگی تقریباً تصادفی مشاهده می‌شود. در شکل (۸) ماتریس‌های درهم‌ریختگی مربوط به مدل‌های مختلف نمایش داده شده‌اند 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 ROC مدل‌های مورد بررسی نشان می‌دهد که میزان تفکیک‌پذیری میان دو کلاس سالم و معیوب در مدل‌های مختلف متفاوت بوده است. همان‌گونه که مقادیر AUC بیان می‌کنند، مدل‌های MobileNet و Custom_CNN با مقادیر ۹۷/۳۱ و ۹۶/۷۵ 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 ResNet50 با مقدار ۸۱/۲ درصد عملکرد متوسطی داشته و مدل EfficientNetB0 با مقدار بسیار پایین ۲۹/۱۶ درصد نشان داده است که توان تمایز بسیار ضعیفی دارد و عملاً خروجی آن نزدیک به حد تصادف است.

شکل منحنی ROC مدل EfficientNetB0 تقریباً 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

بوده و مدل مستعد بیش‌برازش شده است. علاوه بر این، ResNet50 که 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 MobileNet و 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 ۴ ارائه شده است. با وجود آنکه مدل Custom_CNN بیشترین تعداد پارامترها را دارد (حدود ۱۱ میلیون)، مدل MobileNet با تنها حدود ۱۶۴ هزار پارامتر، عملکرد بهتری در تمامی معیارها داشته است. این موضوع نشان می‌دهد که MobileNet در عین سادگی و کم‌حجمی، از لحاظ کارایی محاسباتی بسیار بهینه‌تر است.

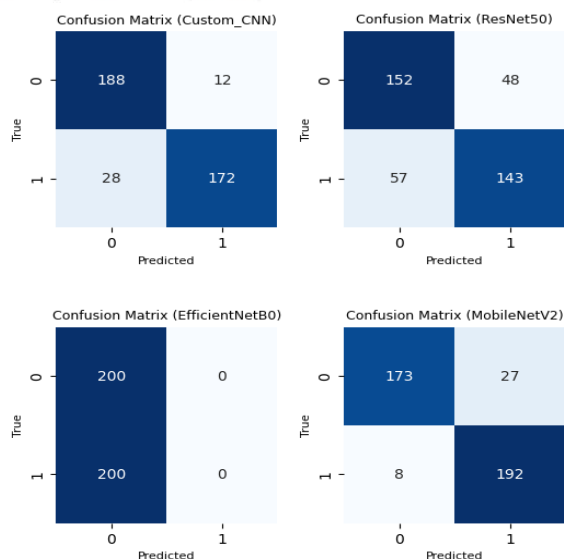
جدول (۴): تعداد پارامترهای قابل آموزش مدل‌های مختلف

مدل	تعداد پارامترهای قابل آموزش
ResNet50	۴۰۱,۲۶۲
EfficientNetB0	۳۴۶,۲۹۰
MobileNetV2	۰۹۷,۱۶۴
Custom_CNN	۰۸۹,۱۶۹,۱۱

۶-۳ تحلیل کیفی و بصری عملکرد مدل‌ها

برای درک بهتر رفتار مدل‌ها، ماتریس‌های درهم‌ریختگی^۱ و منحنی‌های ROC برای تمامی مدل‌ها رسم شده‌اند. ماتریس‌های درهم‌ریختگی نشان می‌دهند که مدل MobileNet و Custom_CNN بیشترین 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 EfficientNetB0 در ماتریس درهم‌ریختگی نشان می‌دهد که این شبکه در تفکیک صحیح نمونه‌های سالم و معیوب

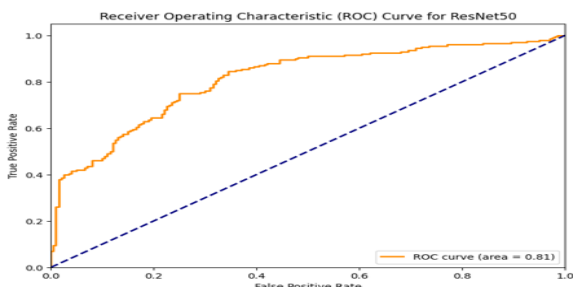
^۱Confusion Matrices



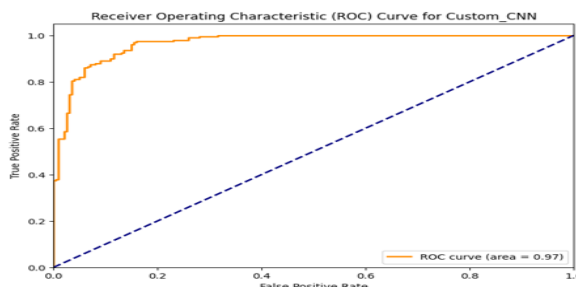
شکل (۸): ماتریس‌های درهم ریختگی مدل‌های مورد بررسی

کلاس در ناحیه‌ای مشابه متمرکز شده و قدرت تفکیک‌پذیری مدل کاهش یافته است. دوم، اشباع تابع فعال‌سازی سیگموئید در لایه خروجی که منجر به تمرکز مقادیر پیش‌بینی در بازه میانی (۰/۴ تا ۰/۶) 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 به صورت 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.

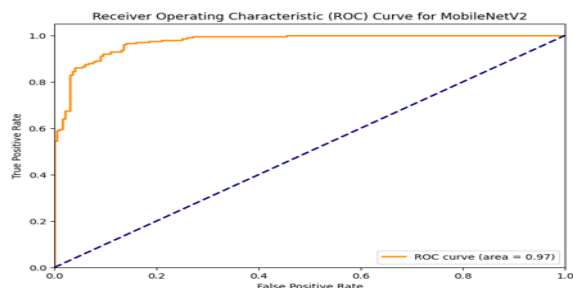
به‌طور کلی، تحلیل منحنی‌های ROC تأیید می‌کند که مدل‌های سبک‌تر و بهینه‌تر از نظر ساختار، در داده‌های صنعتی با نویز و تغییرات نوری بالا، نسبت به مدل‌های عمیق و پیچیده عملکرد دقیق‌تر و پایدارتری دارند.



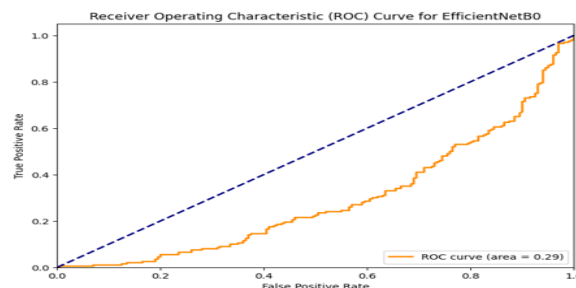
ب) منحنی ROC برای مدل Resnet



الف) منحنی ROC برای مدل پیش‌بینی ساده



د) منحنی ROC برای مدل MobileNet



ج) منحنی ROC برای مدل EfficientNetB0

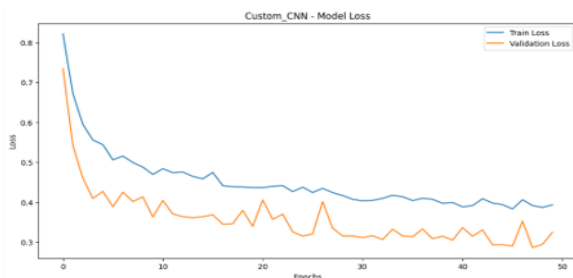
شکل (۹): منحنی‌های ROC مدل‌های مورد بررسی

- 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.

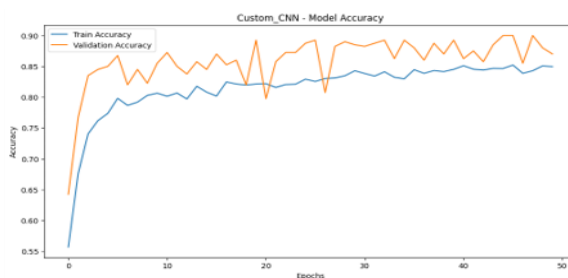
۶-۴ تحلیل فرایند آموزش و همگرایی

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

- مدل‌های ResNet50 و Custom_CNN دارای نوسانات آموزشی قابل توجهی هستند که ناشی از تعداد زیاد پارامترهای قابل یادگیری است.

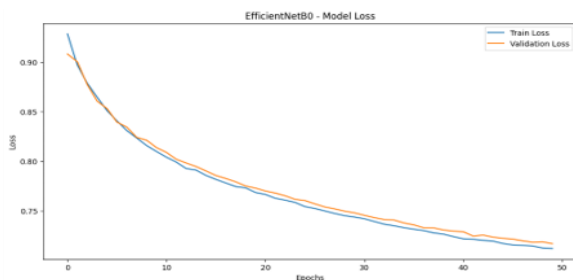


ب) خطا

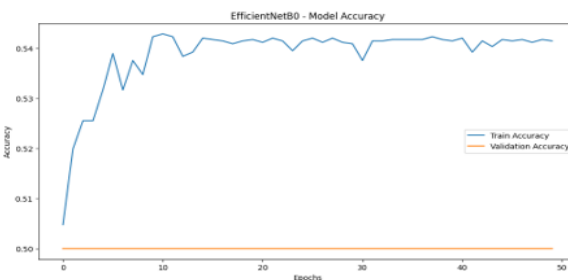


الف) دقت

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

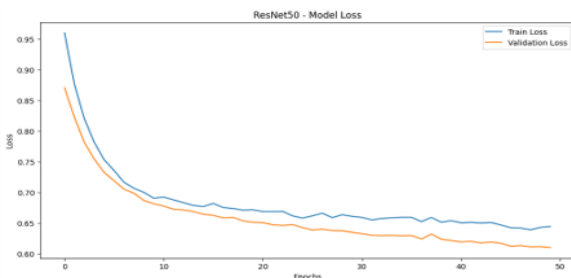


ب) خطا

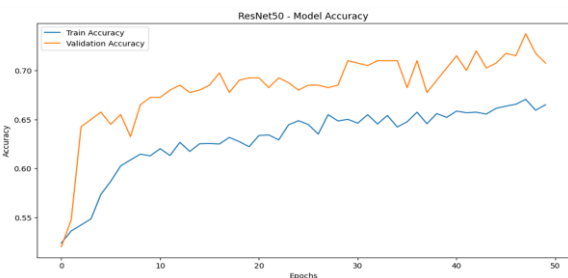


الف) دقت

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

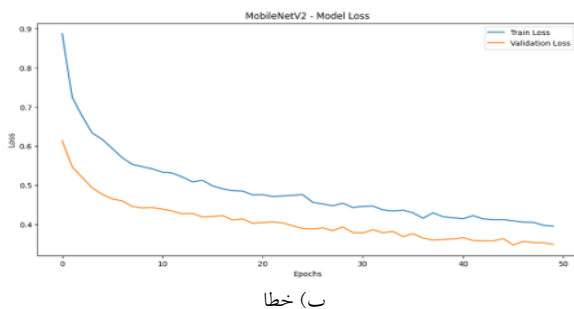


ب) خطا

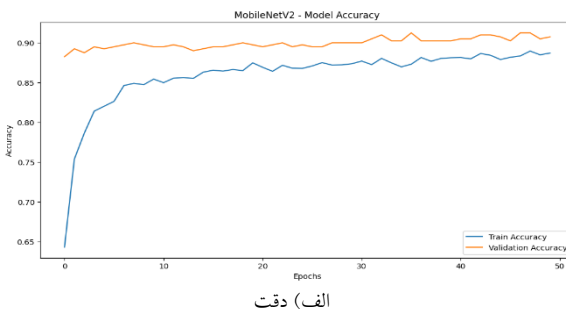


الف) دقت

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

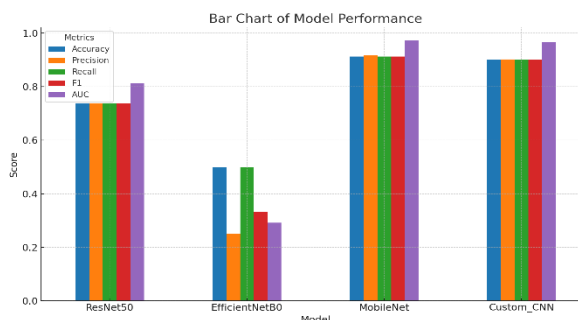
۷- نتیجه‌گیری

در این پژوهش، یک رویکرد جامع و کارآمد برای تشخیص خودکار عیب دوپوستگی در محصولات فولادی ارائه شد. سیستم پیشنهادی با بهره‌گیری از دوربین‌های صنعتی Basler، 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 به منظور 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.

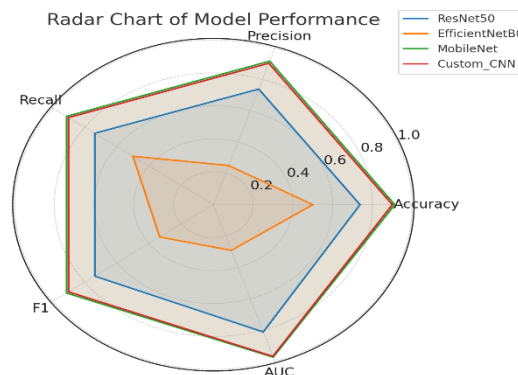
نتایج تجربی نشان دادند که رویکرد پیشنهادی در مقایسه با روش‌های سنتی بازرسی، دقت، سرعت و پایداری بالاتری در تشخیص عیوب سطحی فولاد دارد. به‌ویژه، مدل MobileNet با دقت ۹۱/۲۵ درصد و ساختاری سبک و کارایی محاسباتی بالا توانست بهترین عملکرد را در میان مدل‌های بررسی‌شده، ارائه دهد. استفاده از این سامانه در خطوط 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

در نهایت، به منظور مقایسه جامع عملکرد مدل‌ها، نمودارهای تجمیعی و نمودار راداری^۱ در شکل‌های ۱۴ و ۱۵ ترسیم شده‌اند. این نمودارها نشان می‌دهند که مدل MobileNet در 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

^۱Radar Chart

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.

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

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.

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

Mostafa Majidi^{1*}, Seyedamir Makinejadsanij²

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](#) /© 2025. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

