

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

چکیده

تاریخ دریافت:

۱۴۰۴/۰۶/۱۴

تاریخ پذیرش:

۱۴۰۴/۰۸/۱۶

کلیدواژه‌ها:

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

نویسنده مسئول:

Abolfazl.Gandomi@iau.ac.ir

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 قابل توجه 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 دوبعدی 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 به ترتیب دقت ۹۹/۹۸ و ۹۹/۹۲ درصد به دست آمد که نسبت به مدل پایه VGG16 (حداقل ۴ 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 توسعه یافته و در اختیار متخصصان قرار گرفته است.

doi : 10.22034/ABMIR.2025.23563.1162

E-ISSN: [2821-2037](#)

/© 2023. Published by Yazd University This is an open access article

under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).



۱- مقدمه

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^2 و عدم امکان مدل‌سازی بهینه با اطلاعات مکانی و ویژگی‌های کانالی است. این محدودیت‌ها، می‌تواند منجر به کاهش دقت شود. ازسوی دیگر، 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، مازول توجه و طبقه‌بندی ارائه می‌شود:

مرحله اول؛ استخراج ویژگی با یادگیری انتقال و معماری VGG16^۴.

مرحله دوم؛ ایجاد یک نگاشت توجه فضایی دوبعدی 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 ۴ به شبیه‌سازی و ارزیابی پرداخته می‌شود. در بخش ۵ به ساخت وب اپلیکیشن کاربردی، در بخش ۶ 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 [۴]، در سال ۲۰۱۵ با تمرکز بر طبقه‌بندی اشیاء در زمان واقعی، معماری Faster R-CNN را معرفی کردند. آن‌ها

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 در صورت 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 قابل توجهی 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 معرفی شده‌اند. 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 اشیاء مانند Faster R-CNN و خانواده الگوریتم‌های YOLO^۲ 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

³ Squeeze-And-Excitation

⁴ Visual Geometry Group

¹ Convolutional neural network

² You Only Look Once

مبتنی بر YOLOv5s با ماژول‌های نوآورانه MSFE⁹ (استخراج ویژگی چندمقیاسی) و EFF⁶ (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 BiFPN⁷ 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 دست یافت.

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 دست یافته است. در ارزیابی انجام شده بر روی مجموعه داده NEU-CLS⁸، این مدل به دقت برابر با ۹۹/۴۴ درصد، با SVM و ۹۹/۰۶ درصد با BPNN دست یافت.

وانگ و همکارانش [۱۳] در سال ۲۰۲۵ یک مدل مبتنی بر YOLOv11n با ماژول‌های نوآورانه AMSPPF⁹ (ادغام تطبیقی چندمقیاسی سریع)، C2PSA-DSAM¹⁰ (ماژول توجه فضایی تغییرپذیر)، LDConv¹¹ (کانولوشن خطی تغییرپذیر) و تابع هزینه Inner-CIoU¹² ارائه دادند. در ارزیابی انجام شده بر روی مجموعه داده NEU-DET، این مدل به MAP برابر با ۸۳/۰ درصد دست یافت. لی و همکاران [۱۴] در سال ۲۰۲۵ یک مدل مبتنی بر الگوریتم چندمقیاسی و تقویت ویژگی چندبُعدی ارائه دادند. این مدل از ماژول MRAM¹³ (تجمیع باقی‌مانده چندشاخه‌ای) برای استخراج هم‌زمان ویژگی‌ها و از ماژول MAEH¹⁴ (سر تقویت توجه چندبُعدی) برای ادغام اطلاعات در وجوه مختلف مقیاس و کانال استفاده می‌کند. این مدل بر روی مجموعه داده NEU-DET

جهت ارزیابی از مجموعه داده PASCAL VOC 2007 استفاده کردند و به MAP¹ (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 که توسط خودشان معرفی شد، استفاده کردند و در بخش طبقه‌بندی عیوب به دقت 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 (ASFF³) 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 استخراج 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 به همراه مکانیسم توجه و یک ماژول ادغام چندمقیاسی (MSMP⁵) 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 دست یافت. لی و همکارانش [۱۰] در سال ۲۰۲۴ یک مدل

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

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

۲- ماژول توجه: در این مرحله خروجی مرحله اول شامل ویژگی‌های چندکاناله به شبکه عصبی کانولوشن وارد شده و اعمال زیر به ترتیب 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 به عنوان یک ماسک که به صورت 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^6) 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^۴ 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. بدین ترتیب، حساسیت مدل در تشخیص جزئیات بسیار کوچک افزایش یافته و عملکرد کلی شبکه بهبود قابل توجهی پیدا می‌نماید. در انتهای این بخش یک لایه Dropout جهت جلوگیری از بیش‌برازش 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^۵، شامل ناخالصی^۶، لکه‌ها^۷، سطح حفره‌دار^۸، پوسته نوردی^۹، خراش‌ها^{۱۰} با توجه به نوع خرابی، طبقه‌بندی می‌شوند.

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 انتهای این شبکه، از آن به عنوان استخراج کننده ویژگی‌های اصلی استفاده می‌کند. اما در بخش اصلی مدل از یک شبکه عصبی کانولوشن تغییر یافته با یک ماژول توجه فضایی و یک مکانیزم بزرگ‌نمایی مجدد برای افزایش دقت و پایداری مدل که به اختصار «ماژول توجه» نامیده شده، استفاده شده 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 آموزش دیده، 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.

⁶ Inclusion

⁷ Patches

⁸ Pitted Surface

⁹ Rolled-In Scale

¹⁰ Scratches

¹ Legendre Multiwavelet Transform

² Feature Attention Guidance

³ Global Average Pooling

⁴ Lambda

⁵ Crazing

$C_i(l-1)$ مشخص می‌شود که در اولین لایه کانولوشن برابر با تصویر ورودی ورق‌های فولاد و X_i است.

توابع فعال‌سازی^۲: پس از هر لایه کانولوشن، از یک تابع فعال‌سازی برای غیرخطی بودن مدل استفاده می‌شود. در این تحقیق از توابع زیر استفاده شده است:

تابع فعال‌ساز Sigmoid: مقادیر ورودی را به بازه $[0, 1]$ نگاشت می‌دهد و برای تولید نگاشت توجه ایده‌آل است.

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

خطی^۴: یک تبدیل خطی ساده را بدون اعمال غیرخطی بودن انجام می‌دهد.

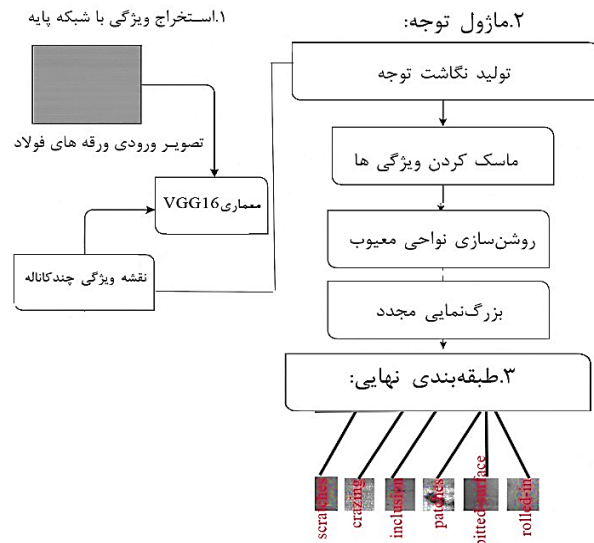
لایه ادغام میانگین^۵: این لایه با محاسبه میانگین مقادیر در یک ناحیه مشخص، ابعاد فضایی نقشه‌های ویژگی را کاهش می‌دهد.

لایه متصل‌کننده کامل^۶: این لایه که در انتهای مدل قرار دارد، وظیفه کلاس‌بندی (طبقه‌بندی) را برعهده دارد.

لایه نرمال‌ساز دسته‌ای^۷: 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^۸ 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^۹ مؤثر، از پدیده بیش برآزش^{۱۰} جلوگیری کرده و قابلیت تعمیم مدل را به شکل قابل توجهی بهبود می‌بخشد [۱۶].

در مدل پیشنهادی، برای بهره‌برداری از مزایای GAP در کنار حفظ حداقلی اطلاعات مکانی، لایه به‌صورت «ادغام میانگین هم‌پوشان^{۱۱}» با گام^{۱۲} برابر با $(1, 1)$ پیاده‌سازی شده است. بررسی مدل با حذف لایه GAP از ساختار شبکه، منجر به نوسانات جزئی در منحنی

در شکل ۱ فلوچارت مدل پیشنهادی نمایش داده شده است.



شکل (۱): فلوچارت مدل پیشنهادی

۲-۳ شرح لایه‌های مدل

لایه کانولوشن^۱: 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:

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} \times C_j^{(l)} \quad (1)$$

در این مدل، $Bi(l)$ ماتریس بایاس در لایه l است. فیلتر یا هسته کانولوشن با نماد $K_{i,j}^{(l-1)}$ نشان داده می‌شود که $m \times m$ اندازه آن است و برای اتصال نگاشت ویژگی شماره j در لایه $l-1$ به نگاشت ویژگی شماره i در لایه l به کار می‌رود. خروجی لایه کانولوشن با نماد $C_i^{(l)}$ نمایش داده می‌شود. ورودی لایه کانولوشن اولیه با نماد

⁷ Batch Normalization

⁸ Spatial Dimension

⁹ Regularizer

¹⁰ Overfitting

¹¹ Overlapping Average Pooling

¹² Stride

¹ Convolutional Layer

² Activation Functions

³ Exponential Linear Unit

⁴ Linear

⁵ Average Pooling Layer

⁶ Fully Connected Layer

جدول (۱): نمادهای محاسباتی مدل پیشنهادی

نماد	توضیح
$F_b n$	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
$F_v gg$	خروجی استخراج ویژگی
$\sigma(\cdot)$	تابع فعال‌ساز Sigmoid
$Conv1 \times 1$	لایه کانولوشن با هسته 1×1 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
$\hat{c}$	۶ کلاس با طبقه‌خرابی پیش‌بینی شده مدل

۴- ماسک‌گذاری ویژگی‌ها در رابطه ۶ بیان شده است:

$$F_{masked} = A \odot F_b n, \quad (6)$$

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

$$F_{masked} = GAP(F_{masked}) \quad (7)$$

$$\bar{A} = GAP(A)$$

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

$$v_{rescaled}(k) = F_{masked}(k) / (\bar{A} + \varepsilon) \quad (8)$$

۷- طبقه‌بندی در رابطه ۹ بیان شده است:

تابع زیان نهایی^۱ 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 $f(0)$ 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 استفاده شده 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.

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

$$I \in \mathbb{R}^{(H \times W \times C)} \quad (2)$$

۲- استخراج ویژگی‌ها با VGG16 و نرمال‌سازی در رابطه ۳ و ۴ بیان شده است:

$$F_v gg = VGG16(I) \quad (3)$$

$$F_b n = BatchNorm(MaxPool(F_v gg)), \quad (4)$$

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

$$A = \sigma(Conv1 \times 1(AvgPool2 \times 2(f(Conv1 \times 1(F_b n)))))) \quad (5)$$

¹ Loss

۳-۴ ساختار الگوریتم مدل پیشنهادی در پایتون

۱- استخراج ویژگی با شبکه پایه:

در ابتدا پیش‌پردازش انجام می‌شود و اندازه تصاویر از سایز $200 \times 200 \times 1$ به سایز $200 \times 200 \times 3$ 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 $2/359/808$ و استخراج ویژگی‌های شبکه پایه به‌صورت نقشه و ویژگی چندکاناله می‌شود.

۲- مازول توجه:

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

▪ تولید نگاشت توجه

این فرایند شامل چندلایه کانولوشنی متوالی با فیلترهای 1×1 است و هدف آن علاوه بر 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 1×1 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 1×1 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 1×1 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.

$$y = \text{Softmax}(W_2 \cdot \text{ReLU}(W_1(v_{r, \text{escaled}}(k)) + b_1) + b_2) \quad (9)$$

$$\hat{c} = \arg \max_i (y_i)$$

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، استفاده شده است. ویژگی‌های کلی دو مجموعه داده به شرح زیر است:

تعداد تصاویر: ۱۸۰۰ تصویر خاکستری^۱، با ابعاد 200×200 پیکسل طبقه‌ها (۶ نوع عیب):

با توجه به انواع خرابی شامل ترک‌خوردگی سطحی^۲، درون ماندگی^۳، پچ‌ها^۴، سطح حفره‌دار^۵، پوسته نوردی^۶ و خراش‌ها^۷ 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 Validation قرار داد و حاوی یک فایل برچسب‌ها^۸ 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 Validation یک فایل XML وجود دارد. در انتها ۱۴۴۸ تصویر در پوشه Train و ۳۶۰ تصویر در پوشه Validation 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 Validation قرار داد و فقط دارای برچسب کلاس (نوع عیب) بوده و مختصات محل عیب ارائه نمی‌شود.

⁵ Pitted Surface

⁶ Rolled-in Scale

⁷ Scratches

⁸ Annotations

¹ Grayscale

² Cracking

³ Inclusion

⁴ Patches

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	۶۵۶۶۴
	conv2d-2	None, 6, 6,) (32	۴۱۲۸
	conv2d-3	None, 6, 6,) (16	۵۲۸
	AveragePooling2D	None, 6, 6,) (16	۰
	conv2d-4	(None, 6, 6, 1)	۱۷
ماسک	conv2d	None, 6, 6,) (512	۵۱۲
	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)	۶۵/۶۶۴
	dropout_2	(None, 128)	۰
	dense_2	(None, 6)	۷۷۴

لایه متصل‌کننده کامل دوم:

خروجی لایه قبلی جهت کلاس‌بندی به این لایه با ۶ نورون و تابع فعال‌ساز Softmax وارد می‌شود و برحسب تعداد عیوب موجود در مجموعه داده به ۶ کلاس ترک‌خوردگی شبکه‌ای، شامل ناخالصی، لکه‌ها، سطح حفره‌دار، پوسته نوردی و خراش‌ها طبقه‌بندی می‌شوند. در جدول ۲ هر ۵ بخش مدل پیشنهادی شامل

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 علاوه بر کاهش اندازه، فرایند نرمال‌سازی تطبیقی هم انجام شود. در نهایت، یک لایه Dropout به‌منظور جلوگیری از بیش‌برازش^۱ به مدل اضافه می‌شود.

■ طبقه‌بندی نهایی

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

لایه متصل‌کننده کامل اول:

¹ Overfitting

بازخوانی: اشاره به کسری از مواردی که به درستی طبقه‌بندی شده به تعداد کل موارد طبقه‌بندی شده در یک کلاس، از رابطه ۱۲ محاسبه می‌شود:

$$\text{recall} = \frac{TP}{TP+FN} \quad (12)$$

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

$$\text{F1_Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

۴- شبیه‌سازی و ارزیابی

با توجه به توضیحات ارائه شده در بخش‌های قبل، در این مقاله از ۱۴۴۸ تصویر پوشه Train از مجموعه داده NEU-DET و NEU-CLS شامل انواع خرابی‌های سطحی در نوارهای فولادی نورد گرم، استفاده شده است. نتایج با اجرای شبیه‌سازی در محیط اسپایدر در نسخه ۵ و در لپ‌تاپ با پردازنده Corei7 و حافظه تصادفی ۲۰ گیگابایت به دست آمده است. نتایج این شبیه‌سازی مرحله به مرحله به نمایش گذاشته می‌شود:

۴-۱ اجرای آنالیز اکتشافی داده‌ها

در این مرحله آنالیز اکتشافی داده‌ها بر ۱۴۴۸ تصویر پوشه Train اجرا شده است. در شکل ۲ تعداد ۱۸ تصویر به صورت تصادفی از کلاس‌های (طبقه‌های) مختلف نمایش داده شده است.

۴-۲ اجرای مدل

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

استخراج ویژگی، تولید نگاشت توجه، ماسک، کاهش ابعاد و طبقه‌بندی به همراه نام لایه‌ها، ساینز داده‌ها در لایه‌ها و تعداد پارامترهای حاصل از هر لایه در شبیه‌سازی پیتون 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 F1^۱، بازخوانی^۲، دقت و صحت^۳ استفاده می‌شود. در معیارهای بالا ۴ پارامتر زیر استفاده می‌شود:

TP^۴: پیش‌بینی درست برای نمونه‌های مثبت (مثلاً ورق فولاد خراب را خراب طبقه‌بندی داده).

FP^۵: پیش‌بینی اشتباه که مدل نمونه منفی را مثبت طبقه‌بندی داده (مثلاً ورق فولاد سالم را خراب طبقه‌بندی داده).

FN^۶: پیش‌بینی اشتباه که مدل نمونه مثبت را منفی طبقه‌بندی داده (مثلاً ورق فولاد خراب را سالم اعلام کرده).

TN^۷: 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 محاسبه شده با یک مقدار طبیعی یا واقعی است. به عبارت دیگر، این معیار ارزیابی می‌تواند مقدار دقیقی که دقت آن قابل اندازه‌گیری است را اندازه‌گیری کند، از رابطه ۱۰ محاسبه می‌شود:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (10)$$

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

$$\text{precision} = \frac{TP}{TP+FP} \quad (11)$$

⁵ False Positive

⁶ False Negative

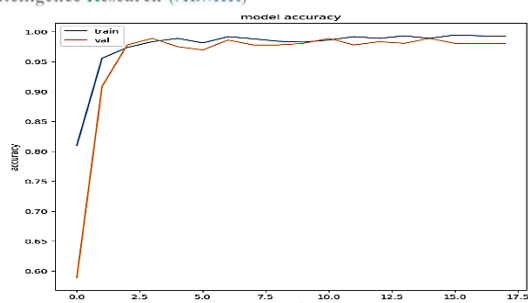
⁷ True Negative

¹ F1-Score

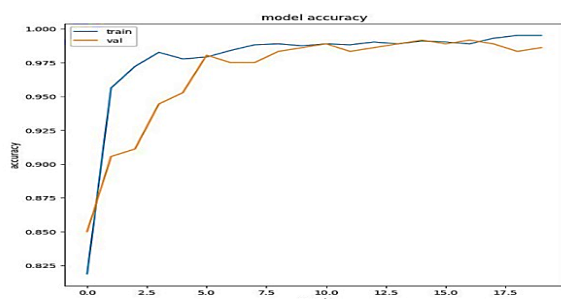
² Recall

³ Precision

⁴ True Positive



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



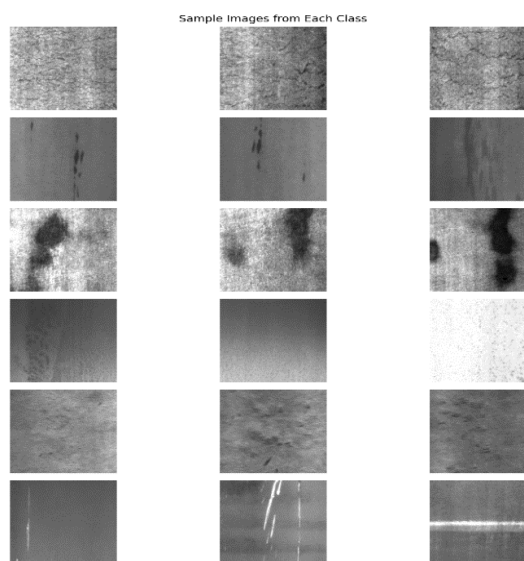
شکل (۴): نمودار دقت در مجموعه داده NEU-CLS

۳-۴ 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 Train و Test اجرا می‌شود. در جدول ۴ بخش‌های مدل در سناریوی ۲ به همراه 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 استفاده شده 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 به ترتیب 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 ۶ نمایش داده شده است:



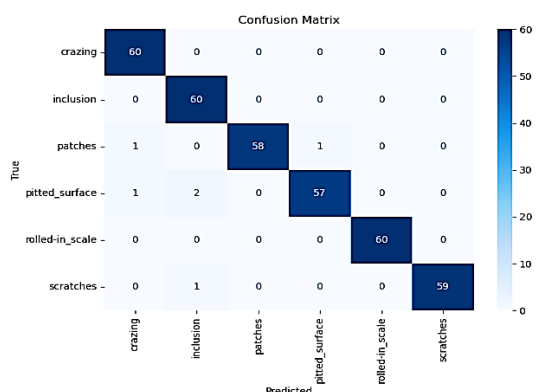
شکل (۲): تعداد ۱۸ تصویر از کلاس‌های مختلف در Train

جدول (۳): ابر پارامترهای مدل پیشنهادی

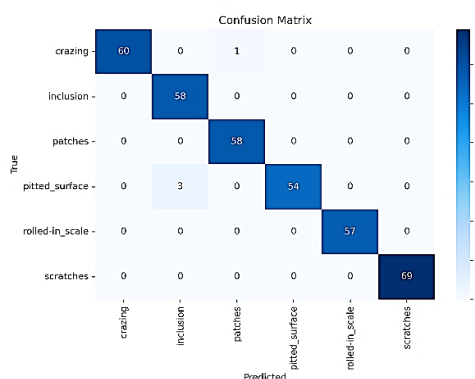
پارامتر	مقدار/انتخاب	توضیحات
ورودی مدل	$3 \times 200 \times 200$	سایز تصویر ورودی
اندازه Batch	۶۴	اندازه بسته‌ها
تعداد epoch	۵۰	تعداد دور آموزش
نرخ Dropout	(۰/۲۵, ۰/۵)	نرخ لایه‌ها
تابع فعال‌ساز	igmoidS, ELU	تابع فعال‌ساز GAP
بهینه‌ساز	Adam	بهینه‌ساز مدل
تابع هزینه	Categorical Cross-Entropy	تابع هزینه برای ۶ کلاس
Callback	Stop_Callback	توقف زودهنگام
Train	۱۱۵۸ تصویر	تعداد تصویر
val	۲۹۰ تصویر	تعداد تصویر
Validation	۳۶۰ تصویر	تعداد تصویر
Learning Rate	۰/۰۰۰۱	نرخ آموزش

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 به ترتیب 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 ۴ نمایش داده شده است:

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



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



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

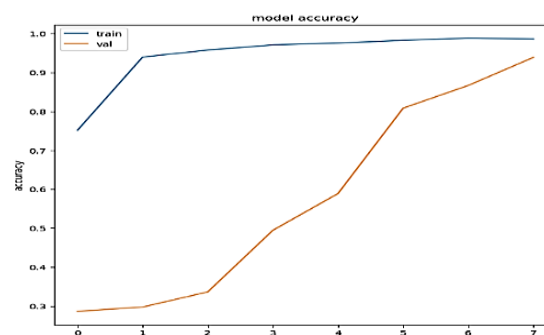
جدول (۵): نتایج سناریوی اول (در قالب درصد)

روش	دقت	دقت مثبت	بازخوانی	امتیاز F1	مجموعه داده
سناریو ۱	۰/۹۴	۹۴/۶۱	۹۳/۸۹	۹۴/۲۱	NEU-DET
سناریو ۱	۹۶/۹۵	۹۵	۹۵/۰۸	۹۴/۹۴	NEU-CLS

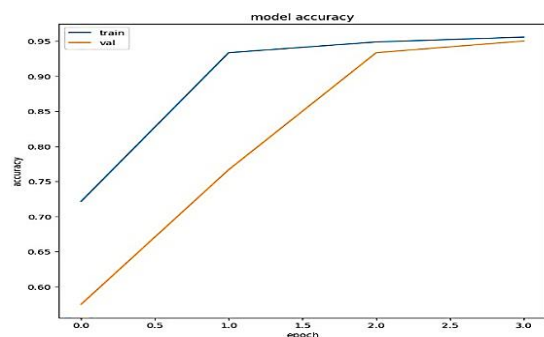
سناریوی دوم: در این سناریو مدل پیشنهادی بر روی ۳۶۰ تصویر به‌عنوان Validation اجرا شده و در شکل ۹ و ۱۰ نتایج حاصل از کلاس‌بندی و پیش‌بینی نوع خرابی در دیتاست NEU-DET و NEU-CLS، در شکل ۱۱ و ۱۲ ماتریس درهم‌ریختگی (در این ماتریس، محور افقی کلاس‌های پیش‌بینی شده و محور عمودی

جدول (۴): ساختار جدولی مدل بدون ماژول توجه در پایتون

پارامترها	سایز داده‌ها	لایه‌ها	مدل
•	(None, ۲۰۰, ۲۰۰, ۳)	input_layer_1	استخراج ویژگی
۱۴/۷۱۵/۷۱۲	(None, ۵۱۲, ۶, ۶)	VGG16--1layer	
•	(None, ۵۱۲)	global_average_pooling2d	
•	(None, ۵۱۲)	dropout_2	
۶۵/۶۶۴	(None, ۱۲۸)	dense_2	طبقه‌بندی
•	(None, ۱۲۸)	dropout_3	
۷۷۴	(None, ۶)	dense_3	

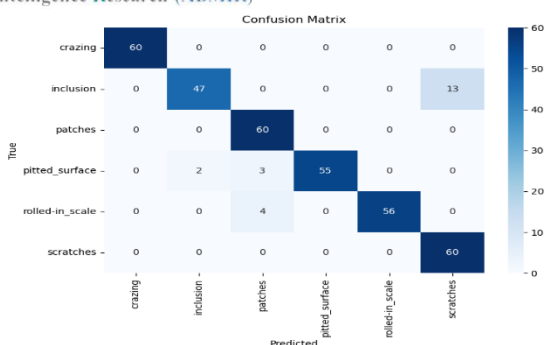


شکل (۵): نمودار دقت مدل در سناریوی ۱ در NEU-DET

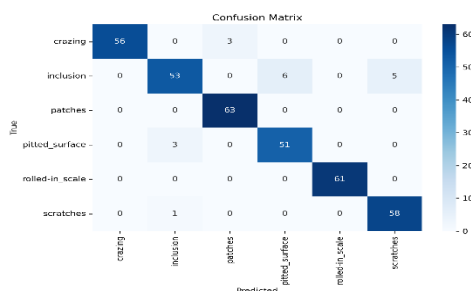


شکل (۶): نمودار دقت مدل در سناریوی ۱ در NEU-CLS

در انتها در شکل ۷ و ۸ ماتریس درهم‌ریختگی (در این ماتریس، محور افقی کلاس‌های پیش‌بینی شده و محور عمودی کلاس‌های واقعی قرار گرفته‌اند. 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:



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



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

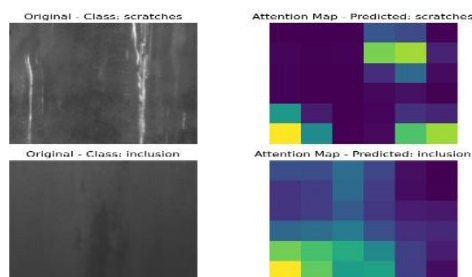
جدول (۶): نتایج سناریو دوم (در قالب درصد)

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

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 ساخته شده است. این ابزار، به کاربران امکان می‌دهد تا با بارگذاری تصاویر ورق‌های فولادی، نوع عیوب موجود را مشخص کند. این اپلیکیشن به صورت خصوصی در [۲۱] دسترس است و در شکل ۱۳، تصویری از رابط کاربری آن نمایش داده شده است.

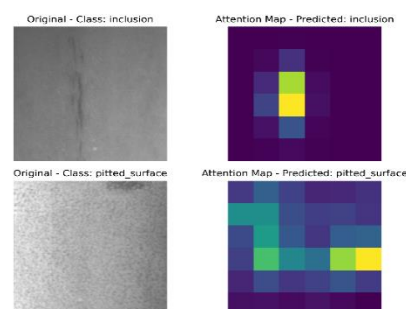
کلاس‌های واقعی قرارگرفته‌اند) در مجموعه داده NEU-DET و مجموعه داده NEU-CLS و جدول ۶ نتایج ارزیابی مدل را در دو سناریو و برای دو مجموعه داده مختلف (NEU-DET و NEU-CLS) نمایش می‌دهد. هر سناریو با مجموعه‌ای از معیارهای ارزیابی اصلی مدل طبقه‌بندی مانند دقت، دقت مثبت، بازخوانی و امتیاز F1 بررسی شده است:



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.

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

سال	محققان	مدل	مجموعه داده	دقت ^۱
۲۰۲۵	سناریو ۱	Attention	N ^۲ -CLS	۹۹/۹۸
۲۰۲۵	سناریو ۲	V ^۳ -Cnn	N-CLS	۹۵/۹۶
۲۰۲۴	ژنگ و همکاران	LWT+ A-E	N-CLS	۹۹/۰۶
۲۰۲۵	ژنگ و همکاران	LWT+FAG	N-CLS	۹۹/۸۰
۲۰۲۵	سناریو ۱	V-Cnn	N-DET	۹۴/۳۸
۲۰۲۵	سناریو ۲	Attention	N-DET	۹۹/۹۲

۱-۶ 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.
- ۴- 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 طراحی شده 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 ۴ شاخص اصلی ارزیابی شامل دقت، پیش‌بینی مثبت، بازخوانی و امتیاز F1 داشته است و همگی به بالاتر از ۹۸٪ رسیده‌اند. نتایج به‌دست‌آمده



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 به همراه 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 به صورت 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

^۳ VGG16

^۱ Percent

^۲ NEU

سطحی ورق‌های فولاد، ضروری است. اگرچه مدل پیشنهادی بر روی دو مجموعه داده معتبر عملکرد مطلوبی نشان داد اما آزمون 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 بازیابی خودکار کنترل کیفیت در صنعت فولاد است. این موضوع نه تنها 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.

Steel Surface Defect Classification Using an Improved VGG16-Based Model

Touba Torabipour¹, Abolfazl Gandomi^{2*}

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

²Assistant 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](#)

/© 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/>).

