



Semnan University



Research Article

DFDA-AD: An Approach with Dual Feature Extraction Architecture and Dual Attention Mechanism for Image Anomaly Detection

Babak Masoudi ^{a,*}

^a Department of Information Technology, Payamenoor University (PNU), P.O.Box, 19395-3697 Tehran, I.R of Iran

PAPER INFO

Paper history:

Received: 2023-07-01

Revised: 2024-04-04

Accepted: 2024-04-07

Keywords:

Abnormality detection;

Attention mechanism;

Deep learning;

Industrial images;

Transfer learning.

ABSTRACT

Detecting and locating unwanted structures or anomalies in the image is one of the important issues in machine vision and industrial inspection. The complexity and variability of data distribution and the lack of labeled data are among the challenges of detecting anomalies in images. In recent years, deep learning methods have provided promising results for solving anomaly detection problems in any data types, especially in images. In this paper, the DFDA-AD architecture, which is an unsupervised approach based on deep learning, is proposed for anomaly detection in industrial images. DFDA-AD consists of dual feature extraction from images by pre-trained DenseNet121 and ResNet50 networks. Two attention mechanisms are improved and developed in this paper, which provide more important feature maps for clustering by K-means algorithm. The evaluation of the model's performance was done on the MVTec AD data set, and the results of the evaluations for anomaly detection and localization were satisfactory compared to several other approaches that have been recently proposed

DOI: <https://doi.org/10.22075/jme.2024.31105.2484>

© 2024 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

* Corresponding Author.

E-mail address: B.masoudi@pnu.ac.ir

How to cite this article:

Masoudi, B. (2024). DFDA-AD: An approach with dual feature extraction architecture and dual attention mechanism for image anomaly detection. *Journal of Modeling in Engineering*, 22(79), 45-57. doi: 10.22075/jme.2024.31105.2484

DFDA-AD: یک رویکرد با معماری استخراج ویژگی دوگانه و سازوکار توجه دوگانه برای تشخیص ناهنجاری در تصاویر

بابک مسعودی^{۱*}

اطلاعات مقاله	چکیده
دریافت مقاله: ۱۴۰۲/۰۴/۱۰	تشخیص و مکان‌یابی ساختارهای ناخواسته و یا ناهنجاری درون تصویر یکی از مسائل مهم در بینایی ماشین و بازرسی صنعتی است. پیچیدگی و متغیر بودن توزیع داده‌ها و نبود داده‌های برچسب‌دار از چالش‌های تشخیص ناهنجاری در تصاویر است. در سال‌های اخیر روش‌های یادگیری عمیق نتایج امیدوارکننده‌ای برای حل مسائل تشخیص ناهنجاری در انواع داده و بخصوص در تصویر ارائه داده‌اند. در این مقاله معماری DFDA-AD که رویکردی بدون نظارت و مبتنی بر یادگیری عمیق است برای تشخیص ناهنجاری در تصاویر صنعتی پیشنهاد شده است. DFDA-AD شامل استخراج دوگانه ویژگی از تصاویر توسط شبکه‌های از پیش آموزش دیده DenseNet121 و ResNet50 است. دو سازوکار توجه در این مقاله بهبود و توسعه داده شده که نقشه‌ی ویژگی‌های مهم‌تر را برای خوشه‌بندی توسط الگوریتم K-means فراهم می‌کنند. ارزیابی عملکرد مدل بر روی مجموعه داده‌ی MVTec AD صورت پذیرفت که نتایج ارزیابی‌ها برای تشخیص و همچنین مکان‌یابی ناهنجاری در مقایسه با چندین رویکرد دیگر که اخیراً پیشنهاد شده‌اند، رضایت بخش بود.
بازنگری مقاله: ۱۴۰۳/۰۱/۱۶	
پذیرش مقاله: ۱۴۰۳/۰۱/۱۹	
واژگان کلیدی:	
تصاویر صنعتی، تشخیص ناهنجاری، ساز و کار توجه، یادگیری انتقالی، یادگیری عمیق.	

DOI: <https://doi.org/10.22075/jme.2024.31105.2484>

© 2024 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

بدون نظارت و یا نیمه نظارت شده انجام شود. روش‌های مبتنی بر تشخیص ناهنجاری با نظارت، داده‌های دارای برچسب را برای آموزش مدل، جهت دسته‌بندی مورد استفاده قرار می‌دهند. در [۲] برای تشخیص خودکار ماژول‌های فتوولتایی^۲ (PV) معیوب در تصاویر مادون قرمز (IR) از داده‌های برچسب‌دار برای آموزش یک شبکه کانولوشنی استفاده شده است. از طرف دیگر در تشخیص ناهنجاری بدون نظارت نیاز به داده‌های برچسب‌دار نیست و بجای آن بر شناسایی الگوهایی که از حالت عادی منحرف شده‌اند، تمرکز شده است. پینایا و همکاران [۳] برای تشخیص ناهنجاری در تصاویر مغزی MRI روشی بدون نظارت مبتنی بر خودرمزگذارهای متغیر^۳ (VAE) و

۱-مقدمه^۱

شناسایی خودکار نواحی غیر عادی یا ناهنجاری درون تصاویر یکی از چالش‌های مهم بینایی ماشین است که در زمینه‌های مختلفی مانند صنعت، مراقبت‌های پزشکی و امنیت کاربرد دارد. به عنوان مثال در حوزه‌ی صنعت بازرسی بصری توسط انسان مشکلاتی نظیر کمبود نیروی کار و بروز اشتباه بدلیل خستگی و یا خطای تصمیم‌گیری را در پی دارد. انتظار می‌رود با تشخیص خودکار ناهنجاری در تصاویر، این مشکلات کاهش یابد. در سال‌های اخیر مطالعات زیادی برای تشخیص ناهنجاری در تصاویر گوناگون انجام شده است [۱]. تشخیص ناهنجاری در تصاویر می‌تواند بصورت با نظارت،

^۲ Photovoltaic

^۳ Variational autoencoder

* پست الکترونیک نویسنده مسی: b.masoudi@pnu.ac.ir

۱. گروه مهندسی فناوری اطلاعات، دانشگاه پیام نور، تهران، ایران

استناد به این مقاله:

سطوح بافت همگن و غیرمنظم، مؤثر نشان داده شده است. برگامن و همکاران [۹] نیز روشی بدون نظارت با استفاده از خود رمزگذارها پیشنهاد کردند. آنها از یک تابع زیان ادراکی بر اساس شباهت ساختاری استفاده کردند که وابستگی‌های متقابل بین نواحی محلی تصویر را با در نظر گرفتن درخشندگی، کنتراست و اطلاعات ساختاری، به جای مقایسه مقادیر تک پیکسلی بررسی می‌کند. نتایج این تحقیق بر روی مجموعه داده‌هایی از مواد نانوالیافی رضایت بخش گزارش شده است. در [۱۰] AnoGAN، یک شبکه متخاصم مولد کانولوشنی عمیق برای یادگیری تنوع آناتومیکی در تصاویر عادی، همراه با یک طرح امتیازدهی ناهنجاری بر اساس نگاشت از فضای تصویر به فضای پنهان، پیشنهاد شده است. نتایج اعمال این مدل روی تصاویر توموگرافی انسجام نوری شبکه نشان می‌دهد که این رویکرد به درستی تصاویر غیرعادی مانند تصاویر حاوی مایع شبکه یا کانون‌های بازتابی را به درستی شناسایی می‌کند. جای و همکاران [۱۱] یک چارچوب تشخیص ناهنجاری مبتنی بر شبکه متخاصم مولد به نام ADE-GAN پیشنهاد کردند. آنها با در نظر گرفتن اطلاعات الگو و نیز خطاهای بازسازی برای بهبود عملکرد تشخیص ناهنجاری، رویکردی جهت محاسبه امتیاز ناهنجاری ارائه دادند. از آنجایی که روش‌های مبتنی بر بازسازی عموماً با محاسبه اختلاف پیکسلی بین تصویر آزمایش و نسخه بازسازی شده ناهنجاری را تشخیص می‌دهند، به کیفیت تصویر تولید شده مخصوصاً در نواحی لبه و یا بافت حساس هستند و ممکن است در عمل ناهنجاری‌ها را به درستی گزارش نکنند. دسته دوم روش‌های تشخیص ناهنجاری بر اساس یادگیری عمیق روش‌های مبتنی بر نمایش^۳ هستند. در این روش‌ها شناسایی بر اساس استخراج ویژگی‌های رمزنگاری شده تصاویر توسط یک شبکه انجام می‌شود. همچنین روش‌های استدلال آماری [۱۲]، استفاده از فضای متغیر پنهان خودرمزگذار [۱۳] و دسته‌بندی کننده GAN [۱۴] نیز در این دسته قرار می‌گیرند. اخیراً در روش‌های متعددی از شبکه‌های کانولوشنی از قبل آموزش دیده بر روی مجموعه داده‌های خارجی در مقیاس بزرگ، مانند ImageNet [۱۵]، برای استخراج ویژگی‌های تصویر استفاده شده است که نتایج بدست آمده از آنها در تشخیص و نیز قطعه‌بندی

ترنسفورمرهای خودهمبسته^۱ را پیشنهاد کرده‌اند. در روش‌های نیمه نظارت شده ترکیبی از داده‌های برچسب‌دار و بدون برچسب برای بهبود دقت مدل استفاده می‌شوند. در [۴] یک روش نیمه نظارت شده مبتنی بر بافت و شیء برای تشخیص ناهنجاری در تصاویر محصول ارائه شده است. این روش از شبکه خودرمزگذار عصبی کانولوشنی (CNN) از قبل آموزش دیده استفاده می‌کند که با یک لایه تبدیل مبتنی بر آمار ترکیب می‌شود تا ناهنجاری‌ها را از تصویر ورودی حذف کند. سپس نسخه "پاک شده"^۲ تصویر ورودی با خود تصویر ورودی مقایسه می‌شود تا از نظر مکانی ناهنجاری‌ها مشخص شوند.

اخیراً از روش‌های یادگیری عمیق بطور مؤثری برای تشخیص ناهنجاری در تصاویر استفاده شده است [۵-۷]. یادگیری عمیق نوعی یادگیری ماشینی است که از شبکه‌های عصبی مصنوعی برای یادگیری، از حجم زیادی داده استفاده می‌کند. این شبکه‌ها از لایه‌های متعددی از نورون‌ها تشکیل شده‌اند که داده‌های ورودی را برای تولید خروجی پردازش و تبدیل می‌کنند. لایه‌های نزدیک‌تر به ورودی مسئول تشخیص ویژگی‌های ساده هستند درحالی که لایه‌های عمیق‌تر ویژگی‌های پیچیده‌تر را شناسایی می‌کنند. یکی از مزایای کلیدی یادگیری عمیق، توانایی آن در یادگیری خودکار نمایش‌های مفید از داده‌های ورودی بدون نیاز به مهندسی ویژگی دستی است. روش‌های تشخیص ناهنجاری مبتنی بر یادگیری عمیق که در مقالات پیشنهاد شده‌اند عموماً در دو دسته قرار می‌گیرند. دسته اول روش‌های مبتنی بر بازسازی^۲ هستند که از یک مدل مولد برای تشخیص ناهنجاری بر اساس محاسبه خطا بین تصویر ورودی و تصویر بازسازی شده، استفاده می‌کنند. استفاده از مدل‌های مولدی نظیر شبکه‌های متخاصم مولد (GAN) و یا خودرمزگذارها برای تشخیص ناهنجاری در تصویر بر این فرضیه استوار است که این مدل‌ها فقط برای فشرده سازی و بازسازی تصاویر معمولی آموزش دیده‌اند و قادر به تعمیم به نمونه‌های غیرعادی نیستند [۸]. می و همکاران رویکردی با بازسازی وصله‌های تصویر توسط شبکه‌های خودرمزگذار حذف نویز کانولوشنی در سطوح مختلف هرم گاوسی، و ترکیب نتایج کانال‌های متفاوت پیشنهاد دادند. نتایج تجربی روش پیشنهادی آنها بر روی

³ Reconstruction-based

⁴ Representation-based

¹ Autoregressive transformer

² Ground truth

تشخیص ناهنجاری نسبت به موارد مشابه از دست آوردهای این پژوهش است.

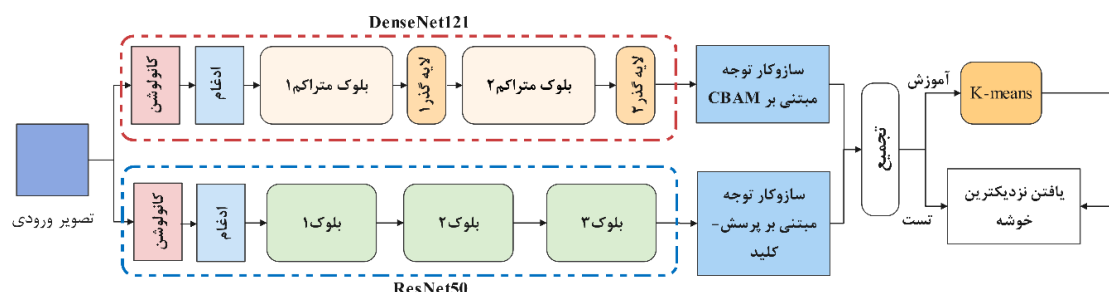
بخش‌های دیگر مقاله به این صورت سازمان‌دهی شده است: در بخش ۲ جزئیات روش پیشنهادی بیان شده است. در بخش ۳ آزمایشات و نتایج و در بخش ۴ جمع‌بندی مطرح شده است.

۲- روش پیشنهادی

شکل (۱) معماری مدل پیشنهادی DFDA-AD برای تشخیص ناهنجاری در تصاویر را نشان می‌دهد. در این مدل برای استخراج ویژگی از تصاویر از دو شبکه از قبل آموزش دیده DenseNet121 و ResNet50 بر روی مجموعه داده ImageNet استفاده شده است. ImageNet یک پایگاه داده‌ی بزرگ از تصاویر دیجیتالی شامل بیش از ۱۴ میلیون تصویر با دقت ۱۰۰۰ پیکسل است که در ۱۰۰۰ دسته مختلف قرار دارند. در معماری پیشنهادی از نقشه ویژگی ایجاد شده در لایه‌های میانی شبکه‌های ذکر شده بعنوان ورودی به دو سازوکار توجه متفاوت استفاده شده است. نقشه ویژگی حاصل از بلوک متراکم^۲ شبکه DenseNet بعنوان ورودی سازوکار توجه CBAM و خروجی بلوک ۳ از شبکه ResNet50 و ورودی سازوکار توجه مبتنی بر پرسش-کلید^۳ است. پس از ادغام خروجی این دو سازوکار توجه به کمک الگوریتم K-means [۲۱] داده‌ها را خوشه‌بندی می‌کنیم. الگوریتم K-means در این مدل علاوه بر اینکه برای خوشه‌بندی مورد استفاده قرار می‌گیرد بدلیل نمونه‌کاهی^۴ داده‌ها انتخاب شده است. در ادامه اجزای مختلف مدل شرح داده می‌شود.

ناهنجاری امیدوارکننده است. در [۱۶] یک رویکرد قطعه-بندی ناهنجاری بر اساس هم‌ترازی بین یک تصویر غیرعادی و تعداد ثابتی از تصاویر عادی مشابه ارائه شده است. در روش تشخیص ناهنجاری هرم معنایی (SPADE) ارائه شده در این مقاله از نقشه‌های ویژگی چند وضوحی و محلی‌سازی بر اساس خوشه‌بندی KNN استفاده شده است. در [۱۷] مدل نرمال بودن با برازش گاوسی چند متغیره (MVG) با ویژگی‌های عمیق شبکه‌های آموزش‌دیده شده در ImageNet تنها با استفاده از داده‌های عادی ایجاد می‌شود. متعاقباً با اعمال فاصله ماکسیمیال امتیاز ناهنجاری محاسبه می‌شود. در [۱۸، ۱۹] از نقشه‌های ویژگی استخراج شده در لایه‌های میانی Wide-ResNet50 [۲۰] که در ImageNet آموزش دیده برای تشخیص ناهنجاری در تصاویر استفاده شده است.

در این مقاله روشی بدون نظارت برای تشخیص ناهنجاری در تصاویر صنعتی پیشنهاد شده است. از آنجا که عموماً داده‌های موجود در صنعت بصورت بدون برچسب و یا با برچسب‌های ناقص موجود هستند این روش بدون نظارت امکان بهره بیشتر از داده‌ها را فراهم می‌کند. مدل پیشنهادی شامل استخراج دوگانه ویژگی‌ها و سازوکار توجه دوگانه است از این رو "DFDA-AD" نامگذاری می‌شود. مدل DFDA-AD ویژگی‌های تصاویر را توسط دو شبکه از پیش‌آموزش‌دیده DenseNet121 و ResNet50 استخراج می‌کند و به کمک سازوکارهای توجه نقشه‌های ویژگی حاوی اطلاعات مهم‌تر را برای تشخیص ناهنجاری جهت دسته‌بندی بدون نظارت، فراهم می‌کند. ارائه یک مدل مبتنی بر یادگیری عمیق و سازوکار توجه و نیز بهبود دقت



شکل ۱- معماری مدل پیشنهادی DFDA-AD برای تشخیص ناهنجاری در تصاویر

³ Query-Key
⁴ Subsampling

¹ Dual feature extraction-Dual Attention mechanism
² Dense block

از مزایای استفاده از این معماری می‌توان به دقت و کارایی بالای آن از نظر تعداد پارامتر اشاره کرد. با توجه به اتصال متراکم، این شبکه قادر است ویژگی‌های تصویر ورودی را بطور موثری استخراج کند که منجر به بهبود دقت می‌شود. علاوه بر این، پارامترهای کمتری نسبت به سایر مدل‌های یادگیری عمیق با دقت مشابه دارد از این رو آموزش و استفاده از آن کارآمدتر است.

۲-۲- شبکه عصبی کانولوشن ResNet50

معماری شبکه عصبی کانولوشن ResNet50 در سال ۲۰۱۵ توسط هی و همکاران [۱۹] معرفی شد. این شبکه از خانواده ResNet است که دارای اتصالات باقی مانده^۴ بین لایه‌ها هستند. معماری ResNet50 از ۵۰ لایه تشکیل شده است و به ۵ مرحله تقسیم می‌شود. ورودی این شبکه نیز یک تصویر سه کاناله و خروجی یک توزیع احتمال softmax بر روی کلاس‌ها است. اتصالات باقیمانده در ResNet50 برای حل مشکل ناپدید شدن گرادیان در شبکه‌های عصبی عمیق استفاده می‌شوند. با افزودن یک اتصال میانبر که یک یا چند لایه را دور می‌زند، گرادیان را می‌توان به راحتی در سراسر شبکه منتشر کرد که منجر به دقت بیشتر می‌شود. هر مرحله در ResNet50 از چندین بلوک باقیمانده تشکیل شده است که هر بلوک از چندین لایه کانولوشن و یک یا چند اتصال باقیمانده تشکیل شده است. تعداد فیلترها در هر لایه کانولوشن با کاهش ابعاد فضایی افزایش می‌یابد که منجر به استخراج ویژگی‌های پیچیده‌تر و انتزاعی‌تر می‌شود. در معماری پیشنهاد شده این مقاله برای استخراج ویژگی از تصاویر بخشی از شبکه ResNet50 استفاده شده است که مشخصات آن در جدول ۲ درج شده است. دقت، کارایی و استحکام بالا نسبت به بیش برآزش شبکه از دلایل انتخاب این شبکه است. به دلیل وجود اتصالات باقیمانده، این معماری می‌تواند ویژگی‌های تصاویر ورودی را بطور موثری استخراج کند که این موجب بهبود دقت مدل می‌شود.

۲-۱- شبکه عصبی کانولوشن DenseNet121

DenseNet121 یک معماری مبتنی بر شبکه عصبی کانولوشنی است که در سال ۲۰۱۶ توسط Huang و همکاران [۲۲] معرفی شد. این معماری با اتصالات متراکم بین لایه‌ها مشخص می‌شود. معماری DenseNet121 از ۱۲۱ لایه تشکیل شده است و به چهار بلوک متراکم و سه لایه گذر^۱ تقسیم شده است. ورودی شبکه یک تصویر سه کاناله و خروجی یک توزیع احتمال Softmax بر روی کلاس‌ها است. بلوک‌های متراکم در DenseNet121 از چندین لایه تشکیل شده‌اند که در آن هر لایه به لایه‌های دیگر به صورت پیشخور^۲ متصل می‌شود. این اتصال متراکم به کاهش تعداد پارامترها و بهبود انتشار ویژگی در سراسر شبکه کمک می‌کند. لایه‌های گذر در DenseNet121 برای کاهش ابعاد فضایی نقشه‌های ویژگی و کنترل تعداد پارامترهای شبکه استفاده می‌شود. در معماری پیشنهادی این مقاله برای استخراج ویژگی از دو بلوک متراکم ۱ و ۲ به همراه دو لایه گذر استفاده شده است که ویژگی‌های معماری استفاده شده در جدول ۱ ذکر شده است.

جدول ۱- مشخصات بخشی از معماری شبکه DenseNet121 که جهت استخراج استخراج ویژگی در مدل DFDA-AD پیشنهادی مورد استفاده قرار گرفته است.

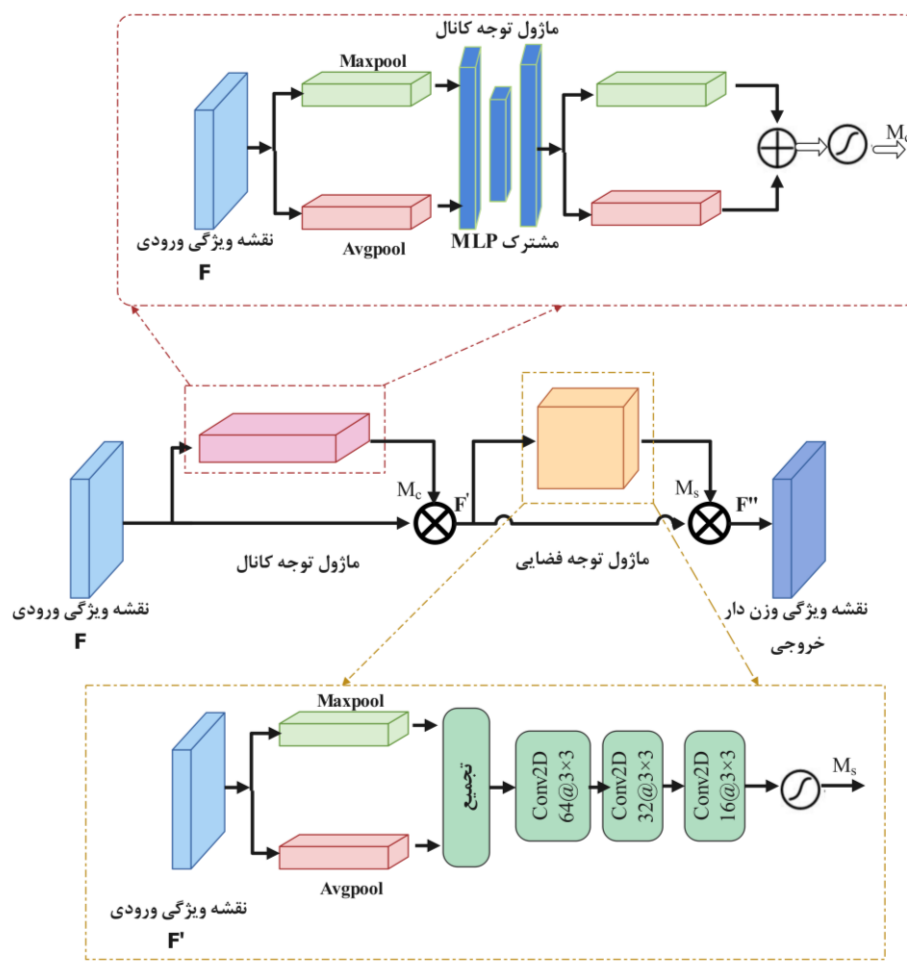
ویژگی معماری	ابعاد خروجی	لایه
$7 \times 7 \text{ Conv, stride } 2$	۱۱۲×۱۱۲	کانولوشن
$3 \times 3 \text{ max pool, stride } 2$	۵۶×۵۶	ادغام ^۳
$[1 \times 1 \text{ conv}] \times 6$ $[3 \times 3 \text{ conv}]$	۵۶×۵۶	بلوک متراکم ۱
$1 \times 1 \text{ Conv}$	۵۶×۵۶	لایه گذر
$2 \times 2 \text{ average pool, stride } 2$	۲۸×۲۸	۱
$[1 \times 1 \text{ conv}] \times 12$ $[3 \times 3 \text{ conv}]$	۲۸×۲۸	بلوک متراکم ۲
$1 \times 1 \text{ Conv}$	۲۸×۲۸	لایه گذر ۲
$2 \times 2 \text{ average pool, stride } 2$	۱۴×۱۴	

^۳ Pooling

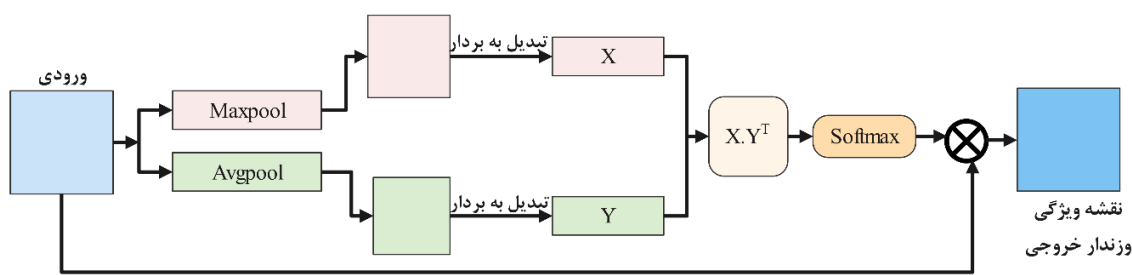
^۴ Residual connections

^۱ Transition layer

^۲ Feedforward



شکل ۲- معماری سازوکار توجه مبتنی بر CBAM



شکل ۳- معماری سازوکار توجه به خود مبتنی بر پرسش-کلید

توجه^۱ CBAM [۲۰] است، را نشان می‌دهد. CBAM نوعی سازوکار توجه است که می‌تواند به شبکه‌های عصبی سیگموئید، اهمیت هر کانال را بصورت یک نقشه ویژگی وزندار محاسبه می‌کند. سپس این نقشه وزندار در نقشه ویژگی اصلی ضرب می‌شود. ماژول توجه فضایی اهمیت هر مکان در نقشه ویژگی‌ها را محاسبه می‌کند. در روش پیشنهادی بجای محاسبه متوالی میانگین و بیشینه ادغام،

۲-۳- ساز و کار توجه مبتنی بر CBAM

شکل (۲) ساز و کار توجه پیشنهادی که مبتنی بر ساز و کار کانولوشنی اضافه شود و عملکرد آنها را در مسائل تشخیص تصویر بهبود دهد. ماژول CBAM از دو بخش تشکیل شده است: ماژول توجه کانال و ماژول توجه فضایی. ماژول توجه کانال با محاسبه میانگین و بیشینه ادغام کانال، عبور دادن آنها از لایه‌های کاملاً متصل و سپس اعمال تابع فعال‌سازی

^۱ Convolutional Block Attention Module

تجمیع با خروجی سازوکار توجه CBAM به 28×28 تغییر پیدا می‌کند.

۳- آزمایشات و نتایج

به منظور ارزیابی کارایی معماری پیشنهادی از تصاویر مجموعه داده MVTEC AD استفاده می‌کنیم. تصاویر مجموعه داده دارای ابعاد 700×700 هستند که به 224×224 تغییر اندازه داده می‌شوند. ۸۰٪ تعداد تصاویر هر زیر مجموعه داده از داده‌های بدون ناهنجاری برای آموزش مدل و مابقی برای ارزیابی در نظر گرفته شد. پیاده سازی مدل توسط پایتورچ و در محیط Google Colab انجام گردید. در فرایند آموزش از تابع بهینه‌ساز ADAM با نرخ یادگیری اولیه 10^{-4} ، اندازه دسته ۱۰ و بیشینه تکرار K-means ۱۰۰ در نظر گرفته شد. برای تجزیه و تحلیل نتایج معیارهای رایج مورد استفاده در تشخیص ناهنجاری که عبارتند از AP^3 ، AP^2 [۲۱]، $AUROC^1$ و [۲۲] محاسبه شد. $AUROC$ معیاری است که توانایی یک الگوریتم تشخیص ناهنجاری را برای تمایز بین نقاط داده عادی و غیرعادی اندازه‌گیری می‌کند. معیار $AUROC$ با ترسیم نرخ مثبت واقعی (TPR) در برابر نرخ مثبت کاذب (FPR) در مقادیر آستانه متفاوت و محاسبه مساحت زیر منحنی حاصل محاسبه می‌شود. مقدار $AUROC$ نشان می‌دهد که الگوریتم قادر است بطور دقیق بین نقاط داده عادی و غیر عادی تمایز قائل شود. $PR-O$ معیار دیگری است که برای ارزیابی عملکرد الگوریتم‌های تشخیص ناهنجاری استفاده می‌شود. این معیار، صحت^۴ و بازیابی^۵ الگوریتم را در مقادیر آستانه متفاوت اندازه‌گیری می‌کند. صحت، نسبت مثبت‌های واقعی را در بین تمام پیش‌بینی‌های مثبت اندازه‌گیری می‌کند در حالی که بازیابی نسبت مثبت‌های واقعی را در بین تمام موارد مثبت واقعی اندازه‌گیری می‌کند. مقدار بالای $PR-O$ نشان می‌دهد الگوریتم قادر است نقاط داده غیر عادی را با دقت بالا شناسایی کند و یادآوری کند. AP معیاری است که مقادیر $PR-O$ را با محاسبه میانگین صحت در سطوح مختلف بازیابی خلاصه می‌کند. مقدار بالای AP نشان دهنده این است که الگوریتم قادر است نقاط داده را با دقت بالا و در مقادیر آستانه مختلف بازیابی کند.

این نقشه‌ها با هم تجمیع شده و بجای دو لایه کاملاً متصل از سه لایه کاملاً متصل عبور کرده و سپس تابع فعال‌سازی سیگموئید به آنها اعمال می‌شود. این نقشه وزندار در انتها با نقشه ویژگی حاصل از مازول توجه کانال ضرب می‌شود. اصولاً مازول توجه CBAM زمانی که در تصاویر الگوها و تغییرات پیچیده وجود دارد میتواند دقت تشخیص را بهبود بخشد.

جدول ۲- مشخصات بخشی از معماری شبکه ResNet50 که جهت استخراج استخراج ویژگی در مدل DFDA-AD پیشنهادی مورد استفاده قرار گرفته است.

لایه	ابعاد خروجی	ویژگی معماری
کانولوشن	112×112	$7 \times 7 \text{ Conv, stride } 2$
ادغام	56×56	$3 \times 3 \text{ max pool, stride } 2$
بلوک ۱	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
بلوک ۲	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
بلوک ۳	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$

۲-۴- سازوکار توجه به خود مبتنی بر پرسش-کلید

سازوکار توجه به خود پیشنهادی مبتنی بر سازوکاری است که در روند Sa-patchcore [۱۴] است. در این سازوکار پیشنهادی علاوه بر ادغام بیشینه که بعنوان پرسش در نظر گرفته شده از ادغام میانگین بعنوان کلید نیز استفاده شده است. شکل (۳) روند این سازوکار را نشان می‌دهد. اگر بردارهای X و Y را به ترتیب مقادیر پرسش و کلید در نظر بگیریم، ماتریس S_{QK} را بصورت زیر تعریف می‌کنیم:

$$S_{QK} = X \cdot Y^T$$

با اعمال تابع فعال‌سازی سیگموئید به S_{QK} نقشه ویژگی وزندار را محاسبه می‌کنیم. این نقشه در ادامه در نقشه ویژگی اصلی ضرب می‌شود. برای کاهش محاسبات در این مقاله رابطه S_{QK} ساده شده است همچنین از خروجی بلوک سوم DenseNet با اندازه 14×14 بعنوان نقشه ویژگی استفاده شده که پس از اعمال سازوکار توجه برای

⁴ Precision

⁵ Recall

¹ Area Under the Receiver Operating Characteristic Curve

² Per Region Overlap

³ Average Precision

۳-۱- معرفی مجموعه داده

مجموعه داده مورد استفاده در آزمایشات این پژوهش مجموعه داده‌ی MVTEC AD [۲۳] است که معیار محبوبی برای تشخیص ناهنجاری در تصاویر بازرسی صنعتی است. این مجموعه داده در سال ۲۰۱۹ توسط برگمان و همکاران معرفی شد و در مجموع شامل ۱۵ دسته شیء مختلف است که هر کدام دارای تعداد متفاوتی از کلاس‌های نقص هستند. در این مجموعه ۵۰۰۰ تصویر در ۱۵ دسته شیء مختلف و ۲۰ تا ۶۰ تصویر در هر دسته وجود دارد. این تصاویر در شرایط نوری مختلف، زوایای متفاوت دوربین و فواصل متغیر تا جسم گرفته شده‌اند. کلاس‌های ناهنجاری در هر دسته بسیار متنوع هستند از خراش و فر رفتگی‌های ساده گرفته تا عیوب پیچیده‌تر مانند تغییر شکل‌ها و تغییر رنگ‌ها. یکی از ویژگی‌های منحصر بفرد مجموعه داده MVTEC AD این است که شامل بافت‌های طبیعی و مصنوعی است که آن را برای الگوریتم‌های تشخیص ناهنجاری چالش برانگیزتر می‌کند. مجموعه داده همچنین شامل تصاویری با وضوح متفاوت است که می‌تواند به ارزیابی استحکام الگوریتم‌های تشخیص ناهنجاری در وضوح متفاوت کمک کند. بطور کلی مجموعه داده MVTEC AD

یک منبع ارزشمند برای محققان و مهندسانی است که بر روی توسعه الگوریتم‌های تشخیص ناهنجاری برای بازرسی صنعتی کار می‌کنند.

۳-۲- ارزیابی تشخیص ناهنجاری

جدول ۳ نتیجه AUROC تشخیص ناهنجاری توسط معماری پیشنهادی را نشان می‌دهد. برای مقایسه بهتر، عملکرد چند مدل دیگر که اخیراً ارائه شده‌اند [۲۴] نیز در جدول آورده شده است. KDAD [۲۵]، CFLOW [۲۶]، DREAM [۲۲] و SSPCAB [۲۷] روش‌هایی مبتنی بر خوشه‌بندی هستند که از خوشه‌بندی k-means برای گروه‌بندی نقاط داده مشابه و شناسایی ناهنجاری به عنوان داده‌هایی که متعلق به هیچ خوشه‌ای نیستند، استفاده می‌کنند. CFA [۲۸] نیز روشی مبتنی بر خوشه‌بندی است که از پوشش محدب^۱ داده‌ها برای تشخیص ناهنجاری استفاده می‌کند. RD4AD [۲۹] از ضرایب وابستگی تصادفی برای اندازه‌گیری شباهت بین نقاط داده و تشخیص ناهنجاری بعنوان نقاط داده با امتیاز RDC کم استفاده می‌کند. PatchCore [۳۰] روشی خوشه‌بندی است که چندین لایه شبکه کانولوشنی و روش نمونه‌کاهی greedy core-set و KNN برای تشخیص ناهنجاری بهره می‌برد.

جدول ۳- نتایج Image-AUROC بر روی مجموعه داده MVTEC-AD

دسته	روش‌ها								
	KDAD	CFLOW	DREAM	SSPCAB	CFA	RD4AD	PatchCore	PRN	روش پیشنهادی
Bottle	۹۹.۲	۹۹.۹	۹۸	۹۵.۶	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۹۹.۶
Cable	۹۰.۳	۹۷.۶	۹۰.۹	۹۲.۷	۹۹.۹	۹۶.۱	۹۹.۹	۹۸.۹	۹۹.۱
Carpet	۸۰.۳	۹۷.۶	۹۶.۹	۹۳.۱	۹۹.۳	۹۸.۷	۹۹.۱	۹۹.۷	۹۸.۹
Capsule	۸۱.۴	۹۷	۹۱.۳	۹۶.۹	۹۷.۴	۹۶.۱	۹۸	۹۸	۹۸.۸
Grid	۷۵.۳	۹۸.۱	۹۹.۹	۹۹.۷	۹۸.۶	۱۰۰	۹۷.۳	۹۹.۴	۹۹.۷
Hazelnut	۹۸.۸	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
Leather	۹۲.۳	۹۹.۹	۱۰۰	۹۸.۷	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
Metal Nut	۷۷.۱	۹۸.۵	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۹۹.۹	۱۰۰	۹۹.۸
Pill	۸۴.۴	۹۶.۲	۹۷.۱	۹۷.۴	۹۷.۷	۹۸.۷	۹۷.۵	۹۹.۳	۹۹.۴
Screw	۸۲.۴	۹۳.۱	۹۸.۷	۹۷.۸	۹۵.۱	۹۷.۸	۹۸.۲	۹۵.۹	۹۸.۳
Transistor	۸۴.۹	۹۲.۹	۹۱.۷	۸۸	۱۰۰	۹۵.۵	۹۹.۹	۹۹.۷	۹۹.۴
Tile	۹۱.۵	۹۷.۱	۱۰۰	۱۰۰	۹۹.۲	۹۹.۷	۹۹.۳	۱۰۰	۱۰۰
Toothbrush	۹۷.۱	۹۸.۸	۱۰۰	۹۷.۹	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
Wood	۹۴.۵	۹۸.۷	۹۹.۵	۹۸.۴	۱۰۰	۹۹.۵	۹۹.۶	۱۰۰	۱۰۰
Zipper	۹۳.۷	۹۷.۱	۱۰۰	۱۰۰	۹۹.۵	۹۷.۹	۹۹.۵	۹۹.۷	۹۹.۸
میانگین	۸۸.۲	۹۷.۵	۹۷.۶	۹۷.۱	۹۹.۱	۹۸.۷	۹۹.۲	۹۹.۴	۹۹.۵

^۱ Convex hull

نسبت به میانگین دقت بهترین الگوریتم مقایسه شده نشان می‌دهد. به این ترتیب نتایج نشان می‌دهند مدل پیشنهادی قادر است بخوبی مکان ناهنجاری را مشخص کند. به منظور مقایسه بهتر دقت مکان‌یابی ناهنجاری جدول ۵ نتایج ارزیابی توسط معیار PR-O و جدول ۶ نتایج ارزیابی توسط معیار AP را نشان می‌دهند. در هر دو معیار ذکر شده میانگین عملکرد مدل پیشنهادی نسبت به بالاترین میانگین سایر روش‌ها بالاتر است. شکل (۴) نمونه‌ای از عملکرد مدل را روی تصاویر مجموعه داده نشان می‌دهد. در این تصاویر مکان ناهنجاری در تصاویر به درستی مشخص شده است.

در انتها PRN [۲۴] از شبکه نمونه اولیه^۱ برای یادگیری رفتار داده‌ها و شناسایی ناهنجاری بعنوان عدم تبعیت از این رفتار استفاده می‌کند. نتایج بدست آمده از این جدول نشان می‌دهد روش پیشنهادی در معیار Image-AUROC عملکرد میانگین بهتری نسبت به سایر روش‌های مقایسه شده از خود نشان داده است.

۳-۳- ارزیابی مکان‌یابی ناهنجاری

جدول ۴ نتایج AUROC برای مکان‌یابی ناهنجاری در تصاویر مجموعه داده را نشان می‌دهد. نتایج pixel-AUROC که برای مکان‌یابی ناهنجاری محاسبه شده است. نتایج جدول ۴ افزایش میانگین دقت را

جدول ۴- نتایج Pixel-AUROC بر روی مجموعه داده MVTEC-AD

دسته	روش‌ها								
	KDAD	CFLOW	DREAM	SSPCAB	CFA	RD4AD	PatchCore	PRN	روش پیشنهادی
Bottle	۹۵.۷	۹۷.۲	۹۹.۱	۹۹.۲	۹۸.۶	۹۸.۸	۹۸.۶	۱۰۰	۱۰۰
Cable	۸۰.۲	۹۷.۸	۹۵.۲	۹۵.۱	۹۸.۸	۹۷.۲	۹۸.۵	۹۸.۸	۹۹
Carpet	۹۵.۵	۹۹.۲	۹۷.۵	۹۲.۶	۹۸.۶	۹۸.۹	۹۹	۹۹	۹۹.۱
Capsule	۹۵.۲	۹۹.۱	۸۸.۱	۹۰.۲	۹۸.۴	۹۸.۷	۹۹	۹۸.۵	۹۸.۷
Grid	۸۹.۴	۹۸.۹	۹۹.۷	۹۹.۵	۹۷.۶	۹۸.۳	۹۸.۷	۹۸.۴	۹۸.۳
Hazelnut	۹۵	۹۸.۸	۹۹.۷	۹۹.۷	۹۸.۶	۹۹	۹۸.۷	۱۰۰	۱۰۰
Leather	۹۸.۱	۹۹.۷	۹۹	۹۶.۳	۹۹.۱	۹۹.۴	۹۹.۳	۱۰۰	۹۹.۸
Metal Nut	۸۳.۳	۹۸.۶	۹۹.۶	۹۹.۴	۹۸.۷	۹۷.۳	۹۸.۳	۹۹.۷	۹۹.۶
Pill	۸۹.۹	۹۸.۹	۹۷.۳	۹۷.۲	۹۸	۹۸.۱	۹۷.۶	۹۹.۵	۹۹.۷
Screw	۹۵.۸	۹۸.۹	۹۹.۳	۹۹	۹۸.۳	۹۹.۷	۹۹.۵	۹۷.۵	۹۸.۸
Transistor	۷۵.۹	۹۸.۲	۸۵.۲	۸۴.۸	۹۸.۱	۹۲.۳	۹۶.۵	۹۸.۴	۹۸.۶
Tile	۸۰.۲	۹۶.۲	۹۹.۲	۹۹.۴	۹۵.۱	۹۵.۷	۹۵.۸	۹۹.۶	۹۹.۸
Toothbrush	۹۵.۵	۹۹	۹۷.۳	۹۷.۳	۹۸.۸	۹۹.۱	۹۸.۶	۹۹.۶	۹۹.۶
Wood	۸۵.۳	۸۶	۹۵.۵	۹۶.۵	۹۴.۷	۹۵.۸	۹۵.۱	۸۲.۹	۹۷.۳
Zipper	۹۵.۳	۹۹.۱	۹۹.۱	۹۸.۴	۹۸.۶	۹۸.۳	۹۸.۹	۹۸.۸	۹۹.۲
میانگین	۹۰	۹۷.۷	۹۶.۷	۹۶.۳	۹۸	۹۷.۸	۹۸.۱	۹۹	۹۹.۱

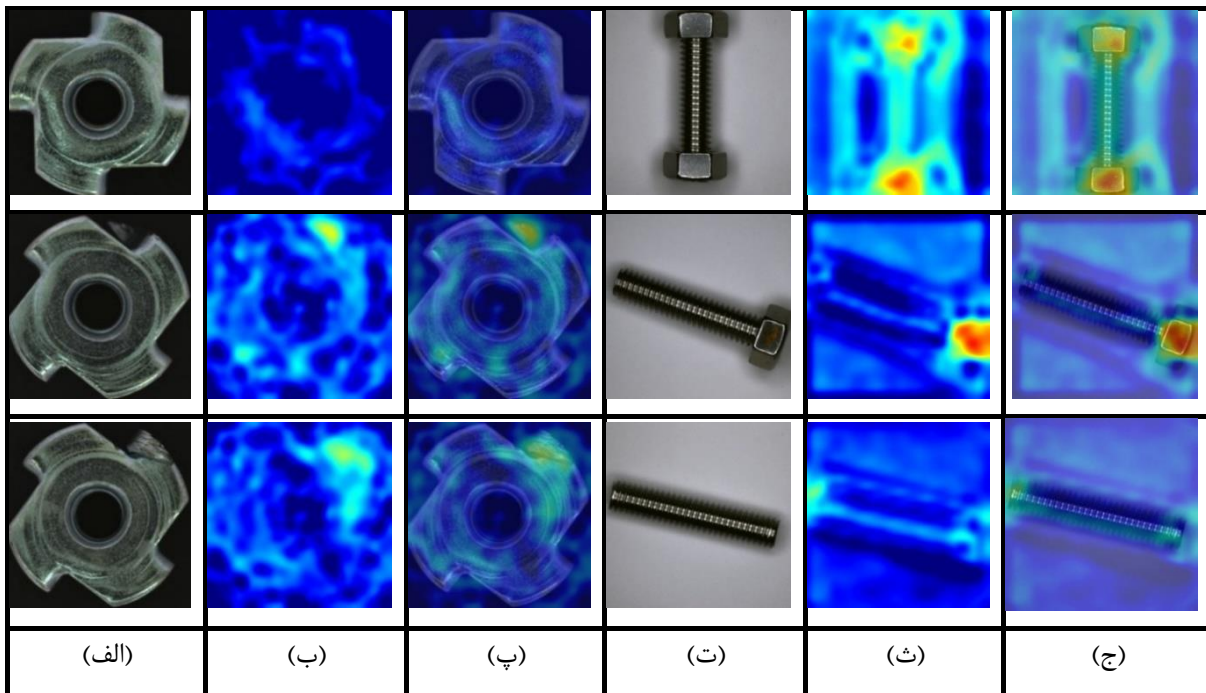
^۱ Prototypical network

جدول ۵- نتایج معیار PR-O بر روی مجموعه داده MVTEC-AD

دسته	روش‌ها								روش پیشنهادی
	KDAD	CFLOW	DREAM	SSPCAB	CFA	RD4AD	PatchCore	PRN	
Bottle	۸۸.۶	۹۴	۹۶.۸	۹۶.۳	۹۴.۶	۹۶.۳	۹۵.۴	۹۷	۹۸
Cable	۶۶.۲	۹۴.۱	۸۱	۸۰.۴	۹۱.۷	۹۴.۱	۹۶.۸	۹۷.۲	۹۸.۱
Carpet	۹۲.۵	۹۷.۶	۹۲.۹	۸۶.۴	۹۳.۶	۹۵.۴	۹۵.۵	۹۷	۹۷
Capsule	۹۰.۱	۹۴	۸۲.۷	۹۲.۵	۹۳	۹۵.۵	۹۳.۴	۹۲.۵	۹۵.۷
Grid	۷۲.۹	۹۶	۹۸.۳	۹۸	۹۲.۹	۹۴.۲	۹۴	۹۵.۹	۹۸.۱
Hazelnut	۹۴.۳	۹۷.۱	۹۸.۵	۹۸.۲	۹۵.۲	۹۶.۹	۹۰.۹	۹۷.۴	۹۸.۳
Leather	۹۷.۵	۹۹.۲	۹۷.۴	۹۴	۹۵.۴	۹۸.۲	۹۶.۹	۹۹.۲	۹۹.۲
Metal Nut	۷۶.۹	۹۱.۵	۹۷	۹۷.۷	۹۱.۴	۹۴.۹	۹۲.۶	۹۵.۸	۹۶.۵
Pill	۸۶.۴	۹۵.۲	۸۳.۴	۸۹.۶	۹۵.۴	۹۶.۷	۹۴.۵	۹۷.۲	۹۷.۴
Screw	۸۵.۲	۹۵.۸	۹۵	۹۵.۲	۹۳.۵	۹۸.۵	۹۷.۵	۹۲.۴	۹۸.۷
Transistor	۶۸.۱	۸۲.۵	۷۰.۴	۶۲.۵	۹۵.۱	۸۳.۳	۹۲.۳	۹۴.۸	۹۶
Tile	۷۴.۳	۸۹.۱	۹۸.۲	۹۸.۱	۸۳.۳	۸۵.۶	۹۱.۳	۹۸.۲	۹۸.۴
Toothbrush	۸۷.۳	۹۵.۳	۸۵.۶	۸۵.۵	۸۶.۸	۹۲.۳	۹۴	۹۵.۶	۹۶.۳
Wood	۷۶.۵	۸۲.۸	۹۰.۳	۹۲.۸	۸۵.۹	۹۱.۴	۸۷.۱	۹۵.۹	۹۵.۸
Zipper	۸۶.۵	۹۶.۶	۹۶.۸	۹۵.۲	۹۴.۳	۹۵.۳	۹۶.۱	۹۵.۵	۹۶.۹
میانگین	۸۲.۹	۹۳.۴	۹۱.۳	۹۰.۸	۹۲.۱	۹۳.۹	۹۳.۹	۹۶.۱	۹۷.۴

جدول ۶- نتایج معیار AP بر روی مجموعه داده MVTEC-AD

دسته	روش‌ها								روش پیشنهادی
	KDAD	CFLOW	DREAM	SSPCAB	CFA	RD4AD	PatchCore	PRN	
Bottle	۵۴.۸	۶۸.۱	۸۸.۹	۸۹.۴	۸۰.۳	۷۸	۷۶.۸	۹۲.۳	۹۳.۶
Cable	۱۲.۶	۶۰.۶	۵۶.۴	۵۲	۷۴.۷	۵۲.۶	۶۷	۷۸.۹	۷۹.۳
Carpet	۴۵.۶	۶۸.۳	۶۵.۱	۴۸.۶	۵۷.۲	۵۶.۵	۶۲.۲	۸۲	۸۳.۲
Capsule	۱۰.۱	۴۸.۸	۳۹.۶	۴۶.۴	۴۸.۳	۴۷.۲	۴۶	۶۲.۲	۶۲.۴
Grid	۷.۳	۴۱.۲	۶۲.۸	۵۷.۹	۲۵.۸	۱۵.۸	۲۴.۵	۴۵.۷	۴۷.۱
Hazelnut	۳۴.۲	۵۹.۹	۹۲.۶	۹۳.۴	۶۰	۶۰.۷	۵۳.۲	۹۳.۸	۹۳.۹
Leather	۲۶.۸	۶۴.۵	۷۲.۹	۶۰.۷	۴۸.۵	۴۷.۶	۴۵.۳	۶۹.۷	۶۹.۵
Metal Nut	۳۴.۱	۸۸	۹۷	۹۴.۷	۹۲.۲	۷۸.۶	۸۶.۶	۹۸	۹۸
Pill	۲۰.۹	۸۲	۴۷.۶	۴۸.۳	۸۱.۹	۷۶.۵	۷۵.۷	۹۱.۳	۹۲.۴
Screw	۶.۱	۴۳.۹	۶۶.۵	۶۱.۷	۲۸.۷	۵۲.۱	۳۴.۷	۴۴.۹	۴۴.۶
Transistor	۲۵.۸	۶۷.۵	۳۹	۳۸.۱	۷۶.۲	۵۴.۱	۶۶.۹	۸۵.۶	۸۶.۱
Tile	۲۷.۷	۶۰.۱	۹۵.۲	۹۶.۱	۵۵.۹	۵۴.۱	۵۶.۲	۹۶.۵	۹۷.۲
Toothbrush	۱۸.۳	۴۶.۳	۴۵.۵	۳۹.۳	۵۵.۷	۵۱.۱	۳۷.۹	۷۸.۱	۷۸.۷
Wood	۲۴.۳	۲۹	۷۴.۶	۷۸.۹	۴۹	۴۸.۳	۴۹.۳	۸۲.۶	۸۳.۶
Zipper	۳۱.۵	۶۵.۲	۷۷.۶	۷۶.۴	۶۵.۲	۵۷.۵	۶۲.۳	۷۷.۶	۷۸.۸
میانگین	۲۵.۳۴	۵۹.۶	۶۸.۱	۶۵.۵	۶۰	۵۵.۴	۵۶.۳	۷۸.۶	۷۹.۲



شکل ۴- نمونه‌ای از خروجی مدل DFDA-AD ستون (الف) تصاویر اصلی دارای ناهنجاری از زیر مجموعه داده metal nut را نشان می‌دهد. ستون (ب) نقشه حرارتی ویژگی‌ها و ستون (پ) تلفیق دو عکس که در هر کدام بخوبی مکان ناهنجاری مشخص شده است. به همین ترتیب ستون (ت) تصاویر اصلی دارای ناهنجاری زیر مجموعه داده Screw را نشان می‌دهد. به همین ترتیب در ستون‌های (ث) و (ج) نقشه ویژگی‌های استخراج شده و تلفیق دو عکس نشان داده شده است.

۴- جمع‌بندی

در این مقاله معماری DFDA-AD برای تشخیص ناهنجاری در تصاویر پیشنهاد شده است. تشخیص مکان یابی ناهنجاری در تصاویر از چالش‌های مهم خطوط تولید صنعتی محسوب می‌شود. معماری پیشنهادی روشی بدون نظارت است که با بهره بردن از یادگیری انتقالی و استخراج ویژگی از تصاویر، توسط دو شبکه از پیش آموزش دیده کارایی مطلوبی را از خود نشان داده است. در پژوهش حاضر دو سازوکار متفاوت توجه پیشنهاد شد که ویژگی‌های

مهمتر را برای خوشه‌بندی فراهم می‌کنند. در پژوهش‌های آتی پیشنهاد می‌شود از معماری‌های مشابه و با بهره بردن از سایر شبکه‌های آموزش دیده جهت تشخیص ناهنجاری در تصاویر استفاده شود. همچنین اعمال سازوکار توجه تأثیر بسزایی در کارایی مدل دارد که می‌توان در چندین لایه میانی شبکه‌ها از آنها استفاده کرد.

تعارض منافع

نویسنده اعلام می‌کند که در مورد انتشار این مقاله تعارض منافع وجود ندارد.

مراجع

- [1] J. Liu, G. Xie, J. Wang, S. Li, C. Wang, F. Zheng, and Y. Jin. "Deep industrial image anomaly detection: A survey." *Machine Intelligence Research* 21, no. 1 (2024): 104-135.
- [2] L. Bommers, M. Hoffmann, C. Buerhop-Lutz, T. Pickel, J. Hauch, C. Brabec, A. Maier, and I.M. Peters. "Anomaly detection in IR images of PV modules using supervised contrastive learning." *Progress in Photovoltaics: Research and Applications* 30, no. 6 (2022): 597-614.
- [3] W.H. Pinaya, P.D. Tudosiu, R. Gray, G. Rees, P. Nachev, S. Ourselin, and M.J. Cardoso. "Unsupervised brain imaging 3D anomaly detection and segmentation with transformers." *Medical Image Analysis* 79 (2022): 102475.
- [4] P. Napolitano, F. Piccoli, and R. Schettini. "Semi-supervised anomaly detection for visual quality inspection." *Expert Systems with Applications* 183. (2021): 115275.
- [5] S. Rezaei, N. Masoud, and A. Khojandi. "GAAD: GAN-enabled Autoencoder for Real-Time Sensor Anomaly Detection and Recovery in Autonomous Driving." *IEEE Sensors Journal* 24, no. 7 (2024): 11734-11742.

- [6] J. Miao, H. Tao, H. Xie, J. Sun, and J. Cao. "Reconstruction-based anomaly detection for multivariate time series using contrastive generative adversarial networks". *Information Processing & Management* 61. no. 1 (2024): 103569.
- [7] G. Zhong, F. Liu, J. Jiang, B. Wang, and C.L.P. Chen. "Refining one-class representation: A unified transformer for unsupervised time-series anomaly detection." *Information Sciences* 656. (2024): 119914.
- [8] J. Yang, Y. Shi, and Z. Qi. "Learning deep feature correspondence for unsupervised anomaly detection and segmentation." *Pattern Recognition* 132. (2022): 108874.
- [9] P. Bergmann, S. Löwe, M. Fauser, D. Sattlegger, and C. Steger. "Improving Unsupervised Defect Segmentation by Applying Structural Similarity to Autoencoders." In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, p. 372-380, 2018.
- [10] T. Schlegl, P. Seeböck, S.M. Waldstein, U. Schmidt-Erfurth, and G. Langs. "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery." In *International conference on information processing in medical imaging*, pp. 146-157. Cham: Springer International Publishing, 2017.
- [11] H. Chai, W. Su, S. Tang, Y. Ding, B. Fang, and Q. Liao. "Improving Anomaly Detection with a Self-Supervised Task Based on Generative Adversarial Network." In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 3563-3567, 2022.
- [12] P. Schlachter, Y. Liao, and B. Yang. "One-class feature learning using intra-class splitting." In *2019 27th European Signal Processing Conference (EUSIPCO)*, pp. 1-5. IEEE, 2019.
- [13] D. Abati, A. Porrello, S. Calderara, and R. Cucchiara. "Latent Space Autoregression for Novelty Detection." In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*: IEEE Computer Society, p. 481-490, 2019.
- [14] Z. Zhang, S. Chen, and L. Sun. "P-KDGAN: progressive knowledge distillation with GANs for one-class novelty detection." In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, Yokohama, Yokohama, Japan*, p. 3237-3243, 2021.
- [15] J. Deng, W. Dong, R. Socher, L.J. Li, L. Kai, and F.F. Li. "ImageNet: A large-scale hierarchical image database." In *IEEE Conference on Computer Vision and Pattern Recognition*, p. 248-255, 2009.
- [16] N. Cohen and Y. Hoshen. "Sub-Image Anomaly Detection with Deep Pyramid Correspondences." *CoRR abs/*(2020): 1-7.
- [17] O. Rippel, P. Mertens, and D. Merhof. "Modeling the Distribution of Normal Data in Pre-Trained Deep Features for Anomaly Detection." In *25th International Conference on Pattern Recognition (ICPR)*: IEEE Computer Society, p. 6726-6733, 2021.
- [18] K. Ishida, Y. Takena, Y. Nota, R. Mochizuki, I. Matsumura, and G. Ohashi. "SA-PatchCore: Anomaly Detection in Dataset With Co-Occurrence Relationships Using Self-Attention." *IEEE Access* 11. (2023): 3232-3240.
- [19] K. Roth, L. Pemula, J. Zepeda, B. Schölkopf, T. Brox, and P.V. Gehler. "Towards Total Recall in Industrial Anomaly Detection." *CoRR abs/*(2021):1-7.
- [20] S. Zagoruyko and N. Komodakis. "Wide Residual Networks." *ArXiv. vol. abs/1606.07146* 2016.
- [21] X. Jin and J. Han, "K-Means Clustering." *Encyclopedia of Machine Learning* (2010): 563-564.
- [22] G. Huang, Z. Liu, L.V.D. Maaten, and K.Q. Weinberger. "Densely Connected Convolutional Networks." In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 2261-2269, 2017.
- [23] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition." In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 770-778, 2016.
- [24] S. Woo, J. Park, J.Y. Lee, and I.S. Kweon. "CBAM: Convolutional Block Attention Module." *Cham: Springer International Publishing, in Computer Vision – ECCV*, p. 3-19, 2018.
- [25] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger. "Uninformed Students: Student-Teacher Anomaly Detection With Discriminative Latent Embeddings." *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019): 4182-4191.

- [26] V. Zavrtanik, M. Kristan, and D. Skocaj. "Draem-a discriminatively trained reconstruction embedding for surface anomaly detection." *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8330-8339, 2021.
- [27] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger. "MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection." *In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 9584-9592, 2019.
- [28] H.M. Zhang, Z. Wu, Z. Wang, Z. Chen, and Y. Jiang. "Prototypical Residual Networks for Anomaly Detection and Localization." *ArXiv abs (2022)*: 1-7.
- [29] M. Salehi, N. Sadjadi, S. Baselizadeh, M.H. Rohban, and H.R. Rabiee. "Multiresolution Knowledge Distillation for Anomaly Detection." *In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 14897-14907, 2021.
- [30] D. Gudovskiy, S. Ishizaka, and K. Kozuka. "CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows." *In 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV): IEEE Computer Society*, p. 1819-1828, 2022.
- [31] N.C. Ristea, N. Madan, R.T. Ionescu, K. Nasrollahi, F.S. Khan, T.B. Moeslund, and M. Shah. "Self-supervised predictive convolutional attentive block for anomaly detection." *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13576-13586, 2022.
- [32] S.L. Sungwook Lee, B.C. Song. "CFA: Coupled-hypersphere-based Feature Adaptation for Target-Oriented Anomaly Localization." *IEEE Access* 10. (2022): 78446-78454.
- [33] H. Deng and X. Li. "Anomaly Detection via Reverse Distillation from One-Class Embedding." *In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR): IEEE Computer Society*, p. 9727-9736, 2022.
- [34] K. Roth, L. Pemula, J. Zepeda, B. Schölkopf, T. Brox, and P. Gehler. "Towards Total Recall in Industrial Anomaly Detection." *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 14298-14308, 2021.