Received: 31 May 2025 | Revised: 26 September 2025 | Accepted: 22 December 2025 | Published online: 31 December 2025

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–16
DOI: 10.47852/bonviewAIA52026321

BON VIEW PUBLISHING

Anomaly Detection Utilizing PatchCore for Reimagining Industrial Visual Inspection

Shalini Kumari[1] **and Chander Prabha**[1,*]

[1] *Chitkara University Institute of Engineering and Technology, Chitkara University, India*

Abstract: In industrial manufacturing, ensuring quality control is critical to maintaining high standards and operational efficiency. Manual defect detection, however, is often time-consuming, error-prone, and costly, thereby driving the need for automated solutions. This paper investigates a technique for industrial anomaly detection (IAD) by utilizing the state-of-the-art PatchCore algorithm in conjunction with the widely recognized MVTec AD dataset. The dataset consists of 5354 high-resolution color images representing diverse objects and textures, including defect-free samples for training and numerous anomalous instances for testing. With over seventy distinct defect types—such as scratches, dents, contamination, and other structural irregularities—the dataset presents substantial challenges for conventional visual inspection (VI) techniques. The approach integrates convolutional neural networks (CNNs) with patch-based feature extraction methods, enabling PatchCore to accurately identify and localize even subtle anomalies within complex industrial imagery. Experimental results demonstrate that PatchCore significantly enhances detection accuracy, reduces false positives, and streamlines the overall inspection process. These improvements have important implications for operational productivity and quality assurance in various industrial sectors, paving the way for more reliable and cost-effective manufacturing practices.

Keywords: anomaly detection, outlier, industrial anomaly detection, machine learning, deep learning

1. Introduction

In anomaly detection (AD), unusual patterns deviating from expected behavior are called outliers. Anomalies can be categorized into three types: point anomalies, which are individual data points that significantly differ from others; contextual anomalies, which are abnormalities that are specific to a certain context; and collective anomalies, where a group of data instances collectively indicates an anomaly. AD is often associated with terms like outlier detection, forgery detection, or out-of-distribution detection. In practical AD scenarios, outliers may be absent, poorly defined, or only present in limited cases. In industrial anomaly detection (IAD), visual inspection (VI) is becoming increasingly difficult due to the continuously rising standards for quality, and all scenarios can be regarded as a form of quality inspection. Quality inspection use cases involve assessing the condition or state of an object rather than identifying flaws or missing components. One use case exclusively focused on quality inspection is determining the state of woven fabrics or leather quality.

Damage detection, also known as defect detection and VI, involves classifying or detecting at least one form of damage. One instance of use cases for damage identification includes identifying surface flaws in internal combustion engine parts or segmenting various steel surface faults [1]. Crack detection is a specific type of damage detection categorized separately due to its frequent occurrence in the literature. The crack detection use case focuses on categorizing, localizing, or segmenting cracks. The typical application scenario is the upkeep of public structures, such as the repair of pavement cracks or concrete cracks. Another application of VI is verifying the presence or absence of a component and identifying any flaws [2]. The completeness check provides a summary of these application cases. A completeness check involves determining the absence or presence of something. The previous VI use case class, designated as "other," encompasses VI use cases that cannot be immediately observed solely through quality inspection and do not fall under damage detection or completeness check categories.

This research work is organized into the following sections: Section 2 presents the current state of research and insights into various AD datasets, along with different types of AD methods used. The limitations of the existing work, following the research gaps and contribution, are presented at the end of Section 2. Section 3 details the methodology, dataset description, and further implementation details. Section 4 presents in detail the result analysis along with a brief discussion. Section 5 further discusses the challenges of industrial AD and their potential solutions. Finally, Section 6 concludes the paper with future directions of research.

2. Literature Review

Researchers have proposed many image AD techniques. Depending on how they work, both machine learning (ML) and deep learning (DL) are used. The study focuses strongly on the industrial image of AD. In addition to giving academics and practitioners a thorough grasp of the advantages and disadvantages of the different DL algorithms for AD in image data, the summary offers insights into the performance of these approaches.

Hyun et al. [3] proposed the ReConPatch technique for identifying irregularities in industrial manufacturing. It employs contrastive representation learning to provide distinctive features for AD. The process works by using a linear modulation on patch features from a model that has already been trained. This creates a representation focused on the target that is easy to distinguish. To tackle the issue of

*Corresponding author: Chander Prabha, Chitkara University Institute of Engineering and Technology, Chitkara University, India. Email: chander.prabha@chitkara.edu.in

insufficient labeled pairings for contrastive learning (CL), pseudo-labels based on pairwise and contextual similarities are employed. Peng et al. [4] proposed a new error detection method using an unsupervised approach. The algorithm is based on defect map prediction using actual manufacturing problems. The method shows a significant advance in error detection in industrial products. It outperforms other detection methods, improving the area under the receiver operating characteristic (AUROC) metric by 1.1%. This shows its strong ability to combine the 15 bands in the MVTec dataset. Furthermore, the method outperforms the best discriminatively trained reconstruction anomaly embedding model (DRAEM) by 31.5% in the detection analysis of average precision (AP) detection, showing significant improvement in fault detection. This is because the method focuses on measuring the distance between normal and abnormal samples rather than knowing exactly what the error is. This network utilizes data processing techniques such as affine transformation and image enhancement to improve noise immunity and robustness. Additionally, they were trained using video images created as input samples. Jiang et al. [5] proposed a novel ground-truth segmentation method to reduce false positives caused by artificial threshold determination. False Rejection (FR)-PatchCore excels in anomaly identification on MVTec datasets and reaches cutting-edge capability on Multiple Product Defect Detection or Multi-Product Defect (MPDD) datasets (with spatial location disparities). The pre-text task of feature-level registration optimizes memory features by reducing registration loss to an ideal feature representation. Experimental results show that FR-PatchCore efficiently handles "similar category" and "spatial transformation category" data, improving its generalization capabilities. Wen et al. [6] focused on the cross-scale with attention normalizing flow (CSA-Flow) novel that incorporates channel attention (CA) and self-attention (SA) modules to enhance AD in high-speed railway systems. This is particularly useful in complex industrial environments, aiming to reduce the need for manual maintenance of high-speed electric multiple units. Intricate backgrounds and enigmatic subjects pose a challenge in detecting defects in manufacturing environments. The CSA-Flow channel feature extraction module takes features of different sizes by combining pretrained convolutional neural network models with a CA module. It also employs the SA module's broader receptive field to gather contextual information. CSA-Flow was evaluated using the high-resolution Synthetic Aperture Radar (SAR) images dataset (HSRBD), which determined that the method achieved the highest rate of anomaly identification. CSA-Flow does not do pixel segmentation but can detect aberrant areas using anomaly scores. Liu et al. [7] introduced a simple but effective method, called SimpleNet, to detect and identify anomalies without monitoring. SimpleNet includes several basic neural network modules designed for training and use in industrial environments. Despite its simplicity, SimpleNet outperforms previously developed methods in terms of performance and implementation speed at the scale of MVTec Anomaly Detection (MVTec AD). SimpleNet provides a new approach bridging academic research and industrial applications in the field of random detection and localization. Roth et al. [8] presented the PatchCore algorithm as a method for detecting and categorizing odd data during testing, using only nominal samples as a basis of knowledge. PatchCore achieves balance by preserving relevant context during testing using memory banks (MB) that contain locally aware, nominal patch-level feature representations generated from ImageNet pretrained networks. It also minimizes runtime by employing coreset subsampling. The outcome is a cutting-edge cold-start image AD and localization system that performs highly on industrial AD benchmarks while requiring few processing resources. At MVTec, the dataset can obtain an AUROC of over 99% in image AD. Ishida et al. [9] introduced SA-PatchCore as an extension of the existing PatchCore algorithm. SA-PatchCore incorporates a self-attention (SA) module to identify anomalies in co-occurrence connections. This nonlinear transformation module may generate feature maps by analyzing the relationship between features without relying on the linear transformation used in traditional self-attention and its training process. SA-PatchCore mitigates the computational complexity of self-attention by utilizing feature maps that have been compressed using a pretrained CNN in the self-attention module. Additionally, the Co-Occurrence anomaly detection-Screw Dataset (CAD-SD) contains local and co-occurrence anomalies, as no AD dataset includes co-occurrence anomalies. SA-PatchCore demonstrates excellent AD performance on MVTec AD, which consists solely of local anomalies. Furthermore, it obtains the highest AD performance in the CAD-SD.

Tang et al. [10] highlight the disparity that currently exists between academic research and its practical implementation in industrial business. A precise and dependable method called Relation-aware Disentangled Learning (RADL) is designed to detect and pinpoint anomalies to address this discrepancy. RADL is a specialized solution that aims to tackle the difficulties associated with inspecting the quality of industrial products in real-world settings. The effectiveness of RADL is validated by the performance attained on the MVTec AD dataset and real-world manufacturing industry datasets compared to earlier studies. The proposed inspection system can improve the effectiveness of manufacturing industry processes in the real world. Choi et al. [11] proposed ViV-Ano, a model that combines a vision transformer (VT) and a Variational Autoencoder (VAE) for AD. The ViV-Ano model demonstrated superior performance to the present model on a test dataset. Models evaluated on the MVTec AD dataset for IAD achieved comparable or superior performance to the previous model. In manufacturing, AD algorithms utilize spatial information to identify flaws in image data. Image analysis and localization techniques enhance decision-making and improve efficiency.

Reconstruction-based AD produced either equivalent or superior results to current anomaly identification methods. Bozcan et al. [12] proposed an architecture that addressed two fundamental issues in smart manufacturing systems, where robots learn tasks from human experts. It first detects anomalies by scoring each observation during the task's execution. The system helps human specialists record innovative demonstrations by recognizing states that differ considerably from training samples. Avoiding repetitive data collection and ensuring a variety of protests are essential for good learning. It is more economical and versatile than parametric models since new data does not require computationally expensive retraining. Jezek et al. [13] addressed the challenge of difficult parallel work of artificial vertical lines in the production of steel elements, and new data were developed to identify errors. Like other AD industry datasets, these data are intended for use by unsupervised and supervised units. Therefore, it includes images without training anomalies, all without anomalies, and with ensemble testing, and mask pairs showing defective areas. Additionally, the dataset is used to evaluate the performance of existing troubleshooting systems for the problems included in the dataset. Traditional detection methods of AD have proven to be ineffective when applied to MVTec AD data, which is considered the method for detecting industrial problems. The Patch Distribution Modeling Framework Patch Distribution Modeling (PaDiM) Framework [14] method showed significant differences in error detection capabilities between the two datasets. Zhao et al. [15] suggested the Patch support vector data description (SVDD) technique for segmenting and detecting image anomalies. The image at the patch level, in contrast to Deep SVDD, also localizes faults. Further, self-supervised learning enhances detection performance. Consequently, the suggested approach attained cutting-edge results on the MVTec AD dataset. Due to their high dimensionality and structure, images were featured in earlier research before the ensuing downstream tasks. Analysis findings indicate that the nearest neighbor method using a raw patch frequently distinguishes abnormalities reasonably well. Kumari et al. [16] suggested the usefulness of DL-based image

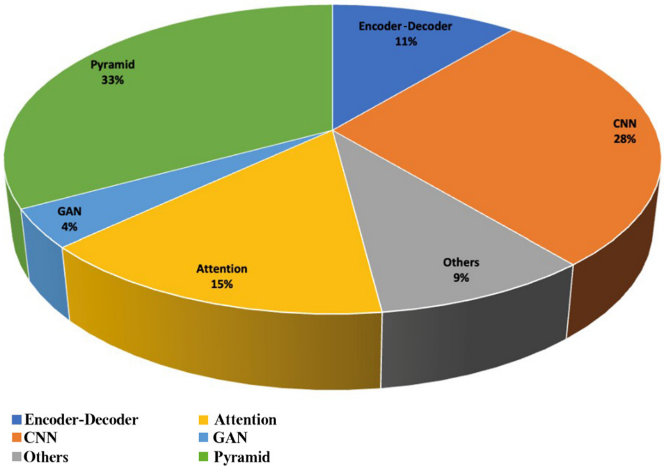
AD in industrial inspection applications. It compares conventional and advanced supervised, unsupervised, and semi-supervised DL approaches in solving important issues like real-time processing, low sample size, and dataset imbalance. The performance is evaluated on various industrial datasets, used in unmanned aerial vehicle (UAV), automated guided vehicle (AGV), and manipulator-based inspection systems, and mitigation strategies and future research directions are discussed to improve AD in manufacturing quality assurance. Kumari et al. [17] discussed the status of vision-based AD in industrial applications in terms of data acquisition, preprocessing, learning mechanisms, and evaluation, categorizing the methods by supervision level: supervised, unsupervised, and semi-supervised learning. Highlighting major challenges, such as real-time processing, small defect detection, and data imbalance, it also shares the solutions, such as edge computing, data augmentation, and generative models for system integration. It also reviews relevant industrial datasets and designs the future directions, such as explainable AI (XAI) and large vision-language models (LVLMs), for improving automated inspection systems. Lin et al. [18] proposed unsupervised industrial image anomaly detection (UIAD) methods in Red Green Blue (RGB), 3D, and multimodal domains. It provides a systematic review of the development of single-modal RGB-based methods, which are mainstream but insufficient for complex scenarios, to the latest 3D methods for spatial data, and multimodal methods for effective detection using fusing RGB, 3D point cloud, and other data. The author classifies some of the state-of-the-art architectures (feature embedding, reconstruction, memory bank, transformer, diffusion, large models), describes some of the major datasets and evaluation metrics, and focuses on multimodal feature fusion approaches. It also presents deployment issues, including a lack of modalities, noise, and domain adaptability, and defines future directions of efficient transferable and noise-resistant algorithms in complex real-world industrial environments.

Liang et al. [19] described Topological/Texture-oriented Contextual Anomaly Detection (ToCoAD) as a two-stage contrastive learning framework for unsupervised AD in images, especially for industrial applications. The first stage uses a discriminative network trained with the help of synthetic anomalies generated with Perlin noise, which allows for the localization of the defects roughly. In the second stage, this network is used to guide a negative bootstrap contrastive learning process that is used to fine-tune the feature extractor and is helpful in overcoming domain gaps between pretrained models and industrial data. Through the use of both synthetic anomalous and positive augmented normal samples, ToCoAD is able to accomplish robust and adaptive feature representations. Experiments on benchmark datasets (MVTec AD, Visual Anomaly (VisA), and BTAD) show competitive and state-of-the-art performance with the pixel-level AUROC above 97%, which shows its effectiveness for industrial defect identification and localization.

Six different DL methods have been used for the detection of defects. These methods include CNN, encoder-decoder, pyramid, generative adversarial networks (GAN), attention mechanism, etc. (Figure 1).

Among all the DL methods used in the literature, the pyramid network approach has the greatest rate at 33%. Because pyramid-based architectures extract features from several layers, improving the results, much recent research has focused on utilizing them for defect detection. Naturally, GAN models account for 4% of DL approaches, whereas encoder-decoder models account for 11%, attention mechanisms for 15%, and other models for 9%.

Figure 1
Comparative analysis of DL methods usage



2.1. Importance of quality control in industry

The industry's ability to produce goods that meet customer expectations depends on quality control. It consists of methodical observation, measurement, and evaluation of every manufacturing stage to find defects and nonconformities. Quality control has several really significant motives. First, it guarantees that objects are safe and not dangerous for users. Second, it encourages consistency and repetition of products, therefore enhancing quality and durability. Last, quality control lowers raw material waste and defects, thus increasing the organization's competitiveness. In today's global corporate climate, any sector seeking profitability and customer satisfaction must also incorporate quality control [20]. Several strategies are applied in production quality control to guarantee product quality and eradicate defects and nonconformities. As a fundamental quality control tool, VI looks for surface, aesthetic, and other flaws in items.

2.2. Artificial intelligence tools in vision systems

Vision systems in many different disciplines employ several artificial intelligence (AI) techniques, for example, Viola–Jones cascade method [21], Scale-Invariant Feature Transform (SIFT) or Speeded-Up Robust Features (SURF) feature descriptors [22], color and texture histograms [23], support vector machines (SVM) [24], decision tree algorithms [25], and, finally, neural networks [26, 27]. Vision systems frequently use neural networks for various important purposes [28]. Table 1 presents the contributions of the researchers' work in vision systems.

2.3. Industrial anomaly detection

Over the past few years, the manufacturing industry has been enhancing the standards for inspecting the quality of industrial products. Precise detection of defects guarantees that industrial products meet the required quality standards and minimizes the risk of safety hazards during their use. The progress of the manufacturing sector has elevated the criteria for assessing the quality of industrial goods [29]. AI-based AD is becoming increasingly important in industrial applications, especially quality control. These methods automatically detect defective items to help manufacturers maintain quality and reduce rework. In the industrial sphere, "anomalies" relate to faults, including scratches, bruises, crushing, foreign colors, and texture alterations. VI inside an industrial setting often reveals these flaws. Product anomalies in industrial settings can significantly impact quality. Image-level and pixel-level detection are two types of AD in computer vision (CV). Image-level detection refers to judging the entire image for anomalies, while pixel-level detection focuses on each pixel. Pixel-level detection yields more accurate and interpretable anomaly

Table 1
Researchers work in vision systems

Year	ML	DL	Contribution	Limitation
2021 [30]	✓	✓	Demonstrated the potential of machine learning and deep learning in the field of operations management.	Vision-based flaw identification is not the primary concern.
2022 [31]	✗	✓	The industrial applications and present techniques of object detecting methods are examined.	Discusses object detection broadly.
2022 [32]	–	–	Introduces unsupervised anomaly localization in industrial images.	Addresses unsupervised machine learning models.
2023 [33]	✗	✓	Evaluations of machine vision applications within the automotive manufacturing sector.	Insufficient emphasis is given on defect detection applications, and learning techniques, along with their evaluation criteria, are not adequately discussed.
2023 [34]	✗	✓	Deep convolutional neural network-based defect detection models in industrial applications.	The study examines techniques based on deep convolutional neural networks (CNNs).
2023 [35]	✗	✓	Exhibits architectures pertinent to deep learning-based intelligent automated decision-making methodologies.	Fails to offer background regarding extensive industrial sectors and their use of IAD algorithms.
2024 [36]	✗	✓	Examines deep learning-based surface defect identification in industrial applications.	Concentrates exclusively on applications based on surface detection.
2024 [37]	✗	✓	Evaluates contemporary approaches in self-supervised AD.	Examines exclusively deep self-supervised learning methodologies. Does not expressly address industrial uses.
2024 [38]	✗	✗	Mapping the implementation of AD algorithms in edge computing.	Examines several articles, a portion of which are not image-based models.
2025 [39]	✗	✓	Techniques employed for the visual examination of air-craft.	Insufficient information is available regarding applications of AI in defect detection.

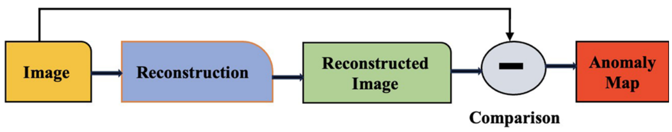
maps. Industrial methods use complete training samples to find faults during testing. It typically targets fine-grained abnormalities in limited image areas, unlike semantic AD, which detects anomalies across the entire image. The main approaches are embedded-based [40–42] and synthesizing-based [43] methods.

For detection, synthesizing-based approaches use augmentations or feature space manipulation [7] to mimic the anomalous distribution. ReContrast [44] and PatchCore [8] emphasize anomalies by embedding normal picture features. These methods are effective at detecting pixel-level anomalies but not semantic ones. For instance, SimpleNet [7], the benchmark's best model, only detects anomalies in local patches. So, it lacks a global perspective needed to model patch cross-correlations, which is crucial for semantic anomaly identification. AD detects anomalies by finding anomalous correlations among patch attributes, not just at the patch level. Over the past decade, AI and deep vision detection technology have advanced in various fields, including autonomous vehicles, surveillance systems, and medical imaging [45, 46]. AI-based deep neural networks (DNNs) are increasingly used in factories to detect product faults because of their higher accuracy and faster inspection speed than previous approaches. AD learning from normal images is a well-liked, unsupervised paradigm for defect detection. The methods are classified as reconstruction based, representation based, and synthesis based.

2.3.1. Reconstruction-based methods

The image reconstruction method [11, 12] is widely used in AD research. The auto-encoder (AE) models the manifold and reconstructs it using the embedding space [13, 47]. The anomalies cannot be rebuilt because they did not evolve during training. The AD result is the difference between the detected and rebuilt images. Various strategies can improve reconstruction results, including GAN [48], learnable memory banks [49, 50], and inpainting masked regions [15]. Although image reconstruction algorithms are practical in industrial settings, they often yield inaccurate findings due to insufficient feature-level discrimination. One drawback of reconstruction-based approaches is their end-to-end learning paradigm, which requires improved network topology, external constraints, and training procedures presented in Figure 2.

Figure 2
Flow of reconstruction-based method

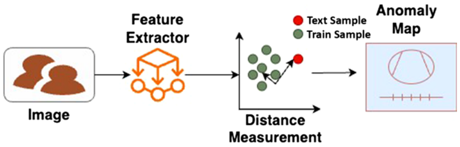


2.3.2. Representation-based methods

AD is a classification problem that differentiates between aberrant and normal data. Tasks that involve categorizing data into two distinct labels are commonly known as binary classifications. Typically, a class label of 0 represents the normal state, whereas a class label of 1 indicates the abnormal condition. During testing, most approaches use the distance between sample features and normal features to find anomalies, such as SPatially-weighted Anomaly Detection (SPADE) [51, 52], PaDiM [52], PatchCore [53], glancing patch (GP) [54], etc., which are common algorithms. These methods use distinct distance measurements (loss functions) to record anomaly scores and build score maps.

The focus your distribution (FYD) [55] approach uses a coarse-to-fine alignment technique to learn about the dense and condensed distribution of regular images. The first alignment stage standardizes the placement of object pixels at both the image and feature levels. Optimal alignment ensures that the features are maximally identical across all positions in the batch. Scale-based image processing approaches include image, patch, and pixel levels. Gaussian-AD utilizes normal images to extract discriminating feature vectors. Patch SVDD [56], PatchCore [53], and PaDiM [52] algorithms utilize normal image patches to generate discriminative feature vectors. SPADE [51, 52] utilizes discriminative features for pixel-level image alignment. These approaches collect normal image features with a statistical method. The premise that anomalous samples have distinct distributions leads to more promising anomaly AD findings in Figure 3.

Figure 3
Flow of the representation-based method



2.3.3. Synthesis-based methods

This method uses nominal (non-defective) images to create fake anomalous images. Using CutPaste [57], false defects (anomalies) are created by randomly pasting a nominal image patch over another nominal image. Synthesizing-based approaches struggle to adequately reflect real anomalies due to their diverse and unexpected appearance. Creating synthetic anomalies from nominal images cannot correctly depict the complexity of real anomalies. Recent OpenGAN [58] research indicates that creating synthesized features, rather than images, improves model performance. This strategy benefits from (1) eliminating noise in feature extraction from synthesized images and (2) reducing model capacity by synthesizing in feature space. To address issues with synthesized images, SimpleNet [59] suggested generating anomalies in the feature space instead of images to address issues with synthesized images.

2.4. Types of datasets and contexts

The nature of data in AD can be categorized along multiple dimensions, such as data type and structure (Figure 4). For the scope of this work, we focus on image-based IAD. Numerous public datasets have been established to benchmark performance in this domain, the most prominent of which are summarized in Table 2.

2.5. Limitations of recent work and research gaps

While the recent literature demonstrates impressive performance on standardized benchmarks, this focus often comes at the expense of critical analysis of broader limitations. Many contemporary studies, including those discussed here, prioritize incremental improvements in metrics like AUROC on datasets such as MVTec AD but often overlook crucial practical deployment challenges. These include (1) extreme computational and memory requirements of methods like PatchCore, which can hinder real-time application; (2) sensitivity to hyperparameters like patch size and coreset ratio, which are rarely subjected to rigorous ablation studies; (3) limited generalization across vastly different domains, as models are typically trained and tested on data from the same distribution; and (4) a lack of robustness to real-world variations like lighting changes, camera angles, and new, unseen defect types that differ from the training set. Truly foundational progress in IAD will require the community to address these practical constraints with the same rigor applied to boosting benchmark scores.

Figure 4
Taxonomy of the dataset

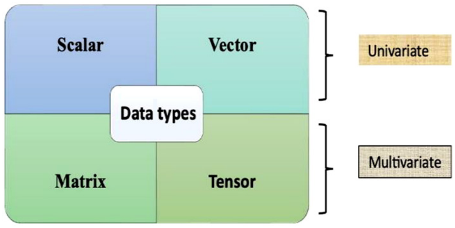


Table 2
Dataset descriptions

Ref.	Dataset	Summary
[6]	Canadian Institute For Advanced Research (CIFAR)	Canadian Institute For Advanced Research (CIFAR)10 and Canadian Institute For Advanced Research (CIFAR)100 have 60,000 natural color images with 32 × 32 resolution. The training set has 50,000 images and the testing set 10,000. Canadian Institute For Advanced Research (CIFAR)10 has 10 equally sized classes, while Canadian Institute For Advanced Research (CIFAR)100 has 100 fine-grained or 20 coarse-grained classes.
[52]	Fashion Modified National Institute of Standards and Technology (MNIST)	Modified National Institute of Standards and Technology (MNIST) has 60,000 training and 10,000 test samples, each made up of 28x28 greyscale images in 10 distinct classes.
[9]	MVTec AD	The MVTec AD dataset, with 5354 high-resolution images from 15 industrial sectors, is a key test for manufacturing AD algorithms. These anomalies have over 70 faults, including scratches, dents, and structural modifications.
[10]	MVTec Local Context (LOCO)	This dataset locates anomalies in images of industrial products with logical and structural faults. It has 3644 industrial inspection- themed images from five classes.
[19]	VisA	Images of manufacturing abnormalities make up the largest industrial anomaly benchmark, Visual Anomaly. It has 10,821 high- resolution images of 12 classes in 3 domains.
[5]	MPDD	This smaller dataset is aimed to detect issues with painted metal part production. The testing setting is realistic, with varied spatial orientations, different items, and diverse backgrounds, unlike lab-based AD datasets. It has 1346 images in 6 categories.

From the limitations above, the research gap is to improve the detection of fine-grained anomalies. To bridge the gap in the literature between theoretical algorithm performance and practical implementation, this paper presents PatchCore and identifies an optimal configuration in terms of backbone, coreset ratio, and feature layers. The usage of the Facebook AI Similarity Search (FAISS) index is a crucial enabler in order to provide key performance and latency numbers necessary to drive adoption in the industry. The contributions of this research work are as follows:

1) A brief comparative analysis of the DL method usage and recent research studies related to image-based IAD, showing existing methodological approaches with utilized datasets.
2) Design and implementation of the proposed technique for IAD using the PatchCore to improve the detection of fine-grained anomalies.
3) The challenges of IAD are identified and examined along with future research directions.

3. Methodology

Anomalib library offers dataset adapters for an increasing number of public benchmark datasets from image domains extensively utilized in the literature. It supports CIFAR-10 [6] for quick prototyping and MVTec [7] and BeanTech Anomaly Detection dataset (BTAD) [8] for real-world defect detection uses. The preprocessing includes changing the input images and dividing the optional image into (non-) overlapping tiles. After that, preprocessing uses the transforms and tiling. In the transforms, Anomalib manages the ground truth pixel mappings with the input images using the Albumentations [11] library for image transformations. Albumentations is a tool for computer vision. Besides its extendable Python interface, Albumentations enables reading transformation parameters from a config file, facilitating experimentation and hyperparameter optimization (HPO). In the tiling, many real-world datasets feature high image resolution, so it is usually necessary to scale the input images before feeding them to the model. Small anomalies in the image could thus lose detail, which makes it more difficult for the model to find these areas. Tiling the input image helps since the size of the abnormal areas stays constant. After that model is deployed along with modular components acting as building blocks to create unique algorithms, Anomalib features a suite of state-of-the-art AD and localization techniques. After that, computational algorithms used in the library are periodically updated with the most recent state-of-the-art AD models. Models now in use might be classified as knowledge distillation models [20], reconstruction [17], and density estimates [12–15].

The model components consist of many ready-to-use modules carrying often-used operations. Like scikit-learn [18], the model components are arranged in relation to their function in AD models (e.g., feature extraction, dimensionality reduction, and statistical modeling). PyTorch implements all model components, enabling all operations to be conducted on the GPU and allowing models to be exported to ONNX and OpenVINO. Applying a custom anomaly detection algorithm using the model components is easy. Like PatchCore [15], consider an anomalous model that first collects features by CNNs and uses coreset sampling [19] to reduce dimensionality. Normalization and thresholding are used in post-processing. In the normalization, depending on the model and dataset, the range of image-level or pixel-level anomaly scores projected by the models in Anomalib during inference could differ. Anomalib normalizes the projected anomaly scores to the [0,1] range, hence transforming the raw anomaly scores into a standardized form. Although Anomalib defaults to min-max normalization with regard to the observed validation values, the normalization technique can be completely turned off or customized. In thresholding, an adaptive thresholding method, which maximizes the threshold value depending on the F1-score during validation, helps the user to select an anomaly score threshold for their trained models. The user may alternatively indicate a manual threshold. In visualization, Anomalib can be set to display and preserve visualizations of the expected anomaly heatmaps and segmentation masks during validation and testing. The PaDiM model is used to discover anomalies. PaDiM uses a novel approach to the cold-start AD method in industrial images. The model, trained on the MVTec AD dataset, effectively localizes industrial abnormalities to their spatial environment without specific dataset training. After teaching the model on normal and defective images, the model detects anomalies by comparing the input image to the learned distribution of normal patches during testing.

3.1. Dataset description

The MVTec AD dataset [59, 60] is a public dataset for industrial inspection. MVTec AD comprises more than 5000 images split into 15 categories. Every category has a training set with nonanomalous samples and a test set that includes both anomalous and nonanomalous samples. Two further datasets include information gathered on separate production lines from several manufacturing sectors. The first dataset consists of medical pills. Single pill images are cropped from a whole image depending on predetermined areas.

Anomaly identification handles every combination of pill blisters. The second dataset of images came from the beverage manufacturing line. While the sizes and colors of restricted caps vary depending on the beverage, an AD system controls all of them. Both datasets maintain the same test and train set structure for every MVTec AD category. The pills dataset's training set comprises 350 RGB images of white oblong tablets set on an aluminum background. Every image has dimensions of $100 \times 210 \times 3$. The test set consists of 224 faulty and 250 non-anomalous samples. Scratches, cracks, incorrect pill orientation, missing sections, or extra pill elements define several flaws in abnormal samples. The dataset of bottle caps comes from Inspect360+ VI systems. One camera gathers cap images, which are then transformed using the Li et al. [61] specified calibration process. The training set contains 243 RGB images with dimensions of $280 \times 96 \times 3$. The test set has 160 normal and 132 aberrant photos with defects: inclined cap, fractured inviolability ring, and absent cap.

3.2. Implementation details

The workflow of developing an IAD system based on the Anomalib library [62] is designed to work with image processing. The pipeline consists of several connected elements structured in a detailed flowchart that executes each task for AD, as presented in Figure 5.

1) Data preprocessing

The preprocessing pipeline was developed to normalize the input images as well as make the model more robust to variations that may occur under industrial conditions. All images were resized to 256×256 pixels, and each image was normalized using ImageNet statistics. Importantly, several normalization and centre-cropping techniques were used during the training on nominal (defect-free) samples to enhance the generalization. This was followed by random rotations (90°, 180°, and 270°), flipping horizontally and vertically, and changes to brightness and contrast. It should be noted that no augmentations that either fake or imitate defects (random erasing, cutout, etc.) were used, as this would poison the nominal training set. The preprocessing steps are presented in Figure 6.

2) Pseudocode for implementation

The implementation of this research is done by extensive custom tuning and tweaking to make it perform the best for the specific industrial contexts under study. This included:

Figure 5
Steps for implementation

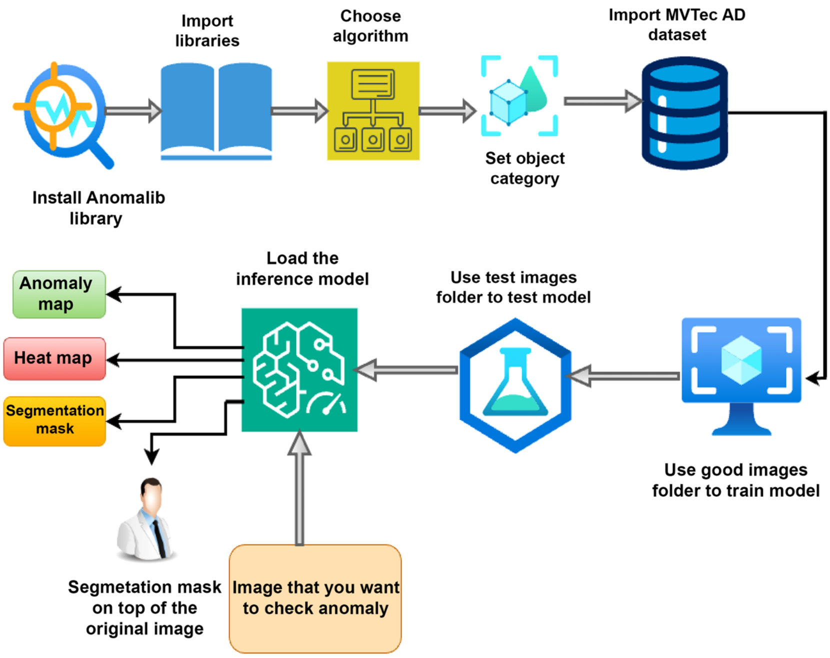
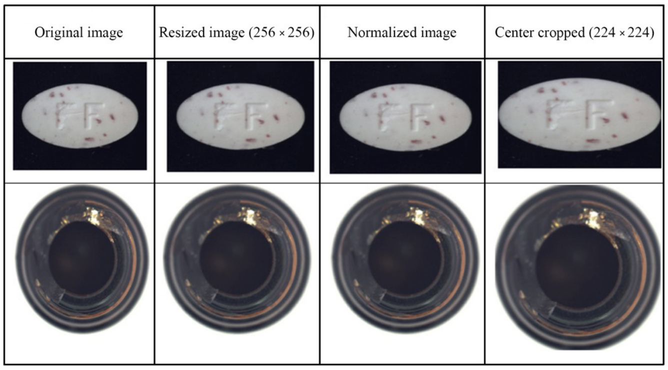


Figure 6
Preprocessing steps of bottle and pills



a. Hyperparameter optimization: The primary parameters of the algorithm, including patch size (3 × 3), coreset sampling ratio (0.25), and DL models for feature extraction (choosing from layer 2 or layer 3 of WideResNet-50-2), were explored using a comprehensive grid search. Table 3 presents the parameter values used for the implementation.
b. Dataset-specific calibration: Anomaly size and contrast are very dynamic factors; therefore, a thresholding scheme was calibrated for each dataset class independently with the help of the F1-maximization concept, departing from the originally adopted global approach.

The PatchCore methodology, illustrated in Figure 7, is engineered for efficiency and effectiveness in a manufacturing context. Its performance stems from a two-stage process:

a. Efficient knowledge representation: Instead of operating on entire images, PatchCore operates on patch-based representations of a pretrained WideResNet-50-2. These features are stored in a "memory bank," which results in a dense representation of normality.

Algorithm Workflow

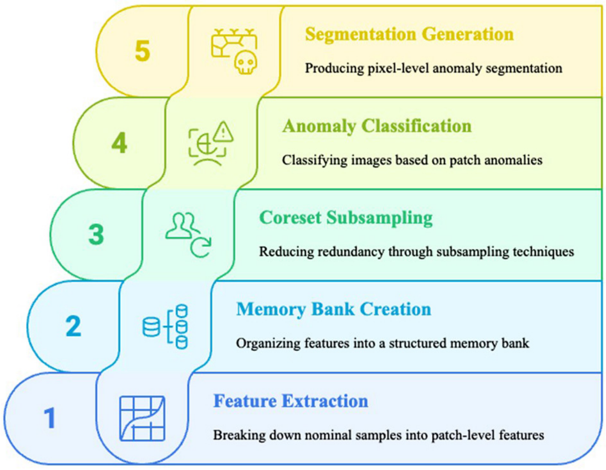
1.	Import and define preprocessing	import anomalib from PIL import Image import torchvision.transforms as transforms preprocess = transforms.Compose([transforms.Resize((256, 256)), transforms.CenterCrop(256), transforms.ToTensor(), transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])])
2.	Load and preprocess image	img = Image.open("your_image.jpg") img_processed = preprocess(img) img_processed = img_processed.unsqueeze(0)
3.	Initialize PatchCore model	from torchvision.models import wide_resnet50_2 from anomalib.models.patchcore import Patchcore backbone = wide_resnet50_2(pretrained=True) model = Patchcore(backbone=backbone)
4.	Train model and build memory bank	for train_imgs in train_loader: features = model.extract_features(train_imgs) model.memory_bank.add(features)
5.	Test model and generate heatmap	features = model.extract_features(img_processed) score, heatmap = model.predict(features, memory_bank=model.memory_bank) show_heatmap(heatmap)

Table 3
Hyperparameter values used for implementation

Parameter	Predefined value	Description
Backbone network	WideResNet-50-2	Pretrained feature extractor (WideResNet-50-2)
Input image size	256 × 256	All images are resized to this resolution during preprocessing
Patch size	3	Size of the local patches extracted from intermediate features
Stride	4	Stride used during patch extraction
Feature layer(s)	layer 2, layer 3	Backbone layers used for feature embedding (usually deeper layers)
Coreset sampling ratio	0.25 (i.e., 25%)	Percentage of patch embeddings used to form the memory bank
Dimensionality reduction	PCA	Absolute necessity for fast NN search
Batch size (inference/train)	32	Batch size used during feature extraction
FAISS index	PQ64 on GPU	Provides the approximate search speed

b. Anomaly as distance: In the process of inference (Figure 7, Steps 4 and 5), the set of anomalies is defined as a simple nearest neighbor search in this patch feature space. The result in the end is an anomaly heatmap with pixel accuracy.

Figure 7
Steps of patch core methodology



PatchCore leverages a pretrained WideResNet-50-2 for feature extraction, generating a rich representation of image patches. To efficiently manage the high dimensionality of these features, a memory bank is constructed. However, storing all features is computationally prohibitive; therefore, a coreset subsampling algorithm is applied to select a maximally representative subset of patches, preserving coverage of the normal data distribution while drastically reducing memory footprint. During inference, the anomaly score for a new image patch is calculated as its distance to the nearest neighbor in the coreset-sampled memory bank.

3) Feature extraction backbone

WideResNet-50-2, a pretrained backbone network on the ImageNet dataset, is used as the PatchCore implementation. This particular model was selected because it has been found to be effective in previous AD research and to provide a good trade-off between representational power and computational efficiency. With the "wide" factor of 2, there are many more filters per layer than in a standard ResNet-50, and so more subtle representations of features important for identifying minor industrial anomalies may be captured. Features were extracted from the predefined intermediate layers of this network (layer 2 and layer 3) in order to extract a hierarchy of information, from fine textures to complex structures, for later processing at the patch level and the construction of memory banks.

The PatchCore analyzes images by using WideResNet-50-2 extracted features to detect normality by assessing feature distribution throughout the feature space. The patch representation features transform the original images into features from another space. PatchCore originally served as an industrial manufacturing AD tool but has no barriers to use in different image fields. This model produces an anomaly heatmap for visual representation. The Anomalib library provides its PatchCore implementation. The initial step of "Feature Extraction" splits nominal samples into their patch-level features at the beginning of this process. The second step in this process moves upward to create "Memory Bank Creation" for structuring collected features into an organized system. The third step utilizes "Coreset Subsampling" procedures to decrease redundancy by applying subsampling methods. "Anomaly Classification" represents Step 4, which involves the image classification process through patch anomaly assessment. After running the process, the last step "Segmentation Generation" creates pixel-level anomaly segmentation, shown in Figure 8. Figure 9 presents memory bank representation with corset sampling, and Figure 10 presents anomaly heatmap of bottle and pill.

4. Results Analysis

The PatchCore was assessed using the standard train-test split provided by the MVTec AD, pill dataset, and bottle cap dataset. The results presented in Table 4 and Figures 11–14 should therefore be interpreted as strong initial indicators of performance on these benchmark datasets, with the understanding that the next critical step is validation under a more rigorous evaluation framework for future work.

For every approach across the three datasets—MVTec AD, pill dataset, and bottle cap dataset—performance measures (AUROC and F1-score). The values show the most successful approaches for every metric and dataset. Some techniques for the bottle cap dataset attained perfect scores (1.0); the table shows the relative performance of several AD techniques that could be readily compared across these significant industrial inspection datasets.

Figure 8
(a) Extracted patches of bottle and (b) extracted patches of pills

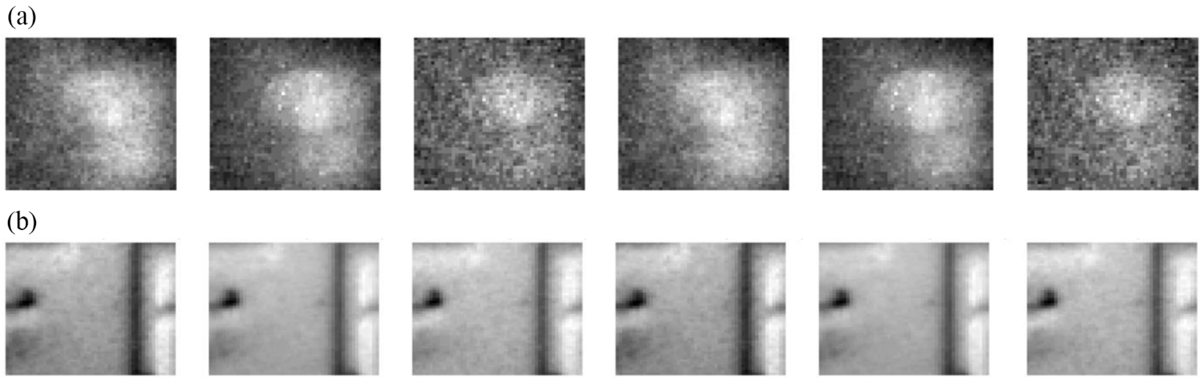


Figure 9
Memory bank representation with corset sampling

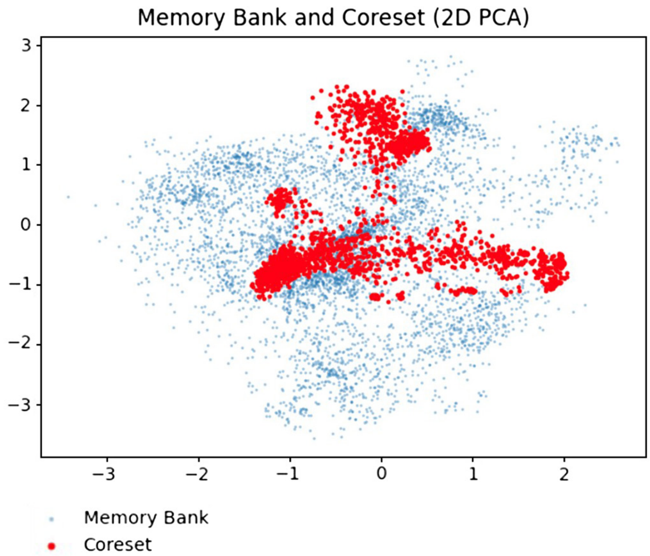
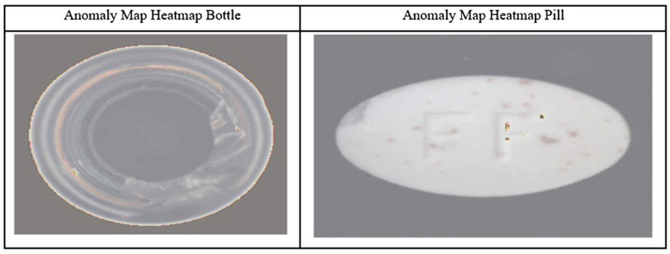


Figure 10
Anomaly heatmap of bottle and pill



AUROC scores performance of six distinct techniques on three datasets—MVTec AD, pill dataset, and bottle cap dataset are shown in Figure 9. AUROC scores falling between 0.70 and 1.05 are shown on the y-axis. With multiple methods obtaining perfect 1.000 scores, the bottle cap dataset usually offers the best performance among approaches. Particularly with DIFAR displaying its lowest score at 0.762, the MVTec AD dataset seems to have the lowest performance for most techniques, with certain methods exhibiting more consistent performance than others.

In addition to the standard AUROC and F1-score (Figures 11 and 12), evaluation of models is done using metrics that are specifically designed for imbalanced data and segmentation tasks. The area under the precision-recall (AUPR) is reported as it is a more reliable performance measure for an AD problem where positive (anomalous) pixels are rare. To measure the spatial precision of anomaly localization, we compute the Intersection over Union (IoU) between the predicted

Table 4
Results comparative analysis

Dataset→	**MVTec AD**		**Pill dataset**		**Bottle cap dataset**	
Metric / **Techniques**	**AUROC**	**F1-score**	**AUROC**	**F1-score**	**AUROC**	**F1-score**
Fastflow	0.979	0.916	0.985	0.957	1.0	1.0
Efficient AD	0.988	0.970	0.961	0.932	1.0	1.0
DFKDE	0.762	0.872	0.954	0.905	0.992	0.969
DFM	0.936	0.943	0.990	0.975	1.0	1.0
PaDiM	0.891	0.916	0.993	0.956	0.996	0.996
PatchCore	0.990	0.976	0.997	0.987	0.986	0.994

segmentation masks and the ground truth. Furthermore, we give the Matthews correlation coefficient (MCC) for a balanced evaluation of the pixel-wise classification performance for all thresholds. The comprehensive metric of MVTec AD is shown in Table 5.

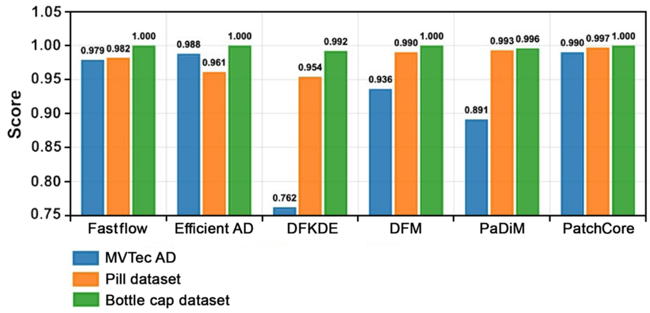
As shown in Table 5, the overall metrics confirm the superior performance of PatchCore. It achieves a very good AUPR score of 0.953, which is an excellent performance on the imbalanced pixel-wise detection task and much better than other models. This means that PatchCore is not only detecting the anomalies, but it does so with high precision and recall.

Table 5
Comprehensive performance metrics on MVTec AD

Model	AUROC	AUPR	F1-score	IoU	MCC
Fastflow	0.979	0.892	0.916	0.843	0.847
Efficient AD	0.988	0.931	0.970	0.912	0.925
DFKDE	0.762	0.455	0.872	0.701	0.723
DFM	0.936	0.823	0.943	0.812	0.835
PaDiM	0.891	0.761	0.916	0.793	0.808
PatchCore	0.990	0.953	0.976	0.931	0.942

The IoU scores give a direct indication of the segmentation quality. PatchCore's IoU of 0.931 shows that there is a very strong spatial overlap between the predictions and the ground truth anomalies, which is critical for accurately localizing faults in an industrial context. The MCC values taking all aspects of the confusion matrix into consideration confirm the overall balanced nature of the performance of PatchCore, a value closer to 1.0 being an almost ideal predictor.

Figure 11
AUROC score comparison



The "AUROC scores" provide a comparison of different techniques. The Score values extend between 0.75 and 1.05 along the vertical axis. Fastflow, Efficient AD, Deep Feature Kernel Density Estimation (DFKDE), DFM, PaDiM, and PatchCore are the six categories on the x-axis. Each category shows three datasets: MVTec AD uses blue, the pill dataset shows orange, and the bottle cap dataset shows green. Each category in the bottle cap dataset leads to the highest recorded scores, which stay near 1.000. The DFKDE category presents the lowest performance for the MVTec AD dataset since it scored 0.762. The pill dataset maintains average performance while exhibiting significant variation when evaluated in the DFKDE category, where it obtained 0.954. A significant finding emerges from PatchCore because all datasets performed well in this aspect, according to the scoring results. Table 6 presents the inference speed benchmarking on MVTec AD.

As you can see in Table 6, PatchCore is a good trade-off between accuracy and speed. While not the fastest model, it has a relatively low inference time of ~200 ms on an edge device, which is fast enough for many industrial inspection applications with cycle times of 0.2 seconds or more (~5 FPS). The main computational bottleneck for PatchCore is the nearest neighbor search between the input patch features and the memory bank that is stored. This is a highly parallelized operation that lends itself well to GPU acceleration, which is confirmed by the high GPU utilization.

The analysis employs the "F1-score" to evaluate six techniques in performance comparison. The score data series extends from 0.850 to 1.050 along the y-axis scale. The chart includes six segments on the x-axis: Fastflow, Efficient AD, and DFKDE, followed by DFM and PaDiM, while PatchCore remains outlined in red. Under the MVTec AD dataset, PatchCore achieved the highest result of 0.976, but DFKDE produced the lowest outcome, 0.872. PatchCore achieved the highest score of 0.987, while DFKDE had the lowest score of 0.905 in evaluating the pill dataset. The average performance of the three datasets is presented in Figure 13.

The PatchCore is shown with a red dashed outline and Efficient AD, DFM, Fastflow, PaDiM, and DFKDE. The presented graph shows two metrics through Score (AUROC & F1) bar charts aligned with both axes. The Score (AUROC & F1) bars span a range from 0.90 to 1.04 on the left y-axis. The right y-axis contains a green line for the Combined Average Score. The methods demonstrate two bars with blue AUROC and orange F1 in a vertical arrangement. Among the methods evaluated, PatchCore delivers the best performance of 0.996 (blue) and 0.988 (orange); Efficient AD, DFM, and Fastflow follow, while DFKDE demonstrates the lowest scores at 0.903 (blue) and 0.915 (orange). The green line within the graph moves from left to right while showing that PatchCore delivers superior results on all metrics through the datasets. Figure 14 presents the AD performance heatmap.

The heatmap under the "AD performance" label compares the six AD techniques (Fastflow, Efficient AD, DFKDE, DFM, PaDiM, and PatchCore) while evaluating their performance on MVTec AD, pill

Table 6
Inference speed benchmarking on MVTec AD

Model	Hardware platform	Avg. inference time (ms)	FPS	GPU utilization (%)	CPU utilization (%)
PatchCore	NVIDIA V100	45	22	85	15
PatchCore	NVIDIA Jetson Xavier AGX	210	4.8	95	40
Efficient AD	NVIDIA V100	12	83	–	–
PaDiM	NVIDIA V100	90	11	–	–

dataset, and bottle cap dataset. The scores of each dataset for AUROC and F1-score metrics are shown through a gradient scale, which transitions from blue (lower scores) to dark red (higher scores). Among all techniques, the bottle cap dataset performs best by achieving a perfect score of 1.000. On MVTec AD, DFKDE demonstrates considerably inferior performance, achieving 0.762 AUROC and 0.872 F1-score. The PatchCore achieves the highest score average of 0.992 across every dataset. Strong results appear in all techniques for the pill dataset, yet performance outcomes for the MVTec AD dataset remain inconsistent. Overall performance statistics show that PatchCore maintains the top position with average results of 0.992, while DFM (0.974), Efficient AD (0.975), and Fastflow (0.973) occupy the following positions.

Figure 12
F1-score comparison

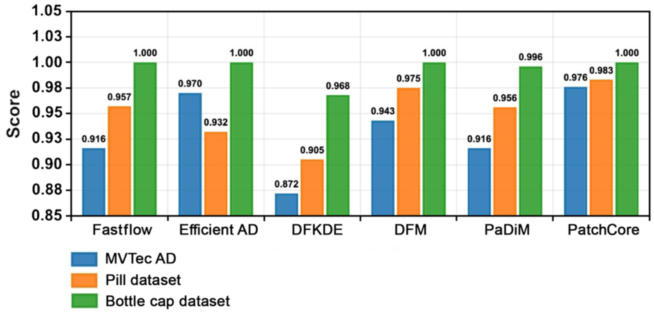
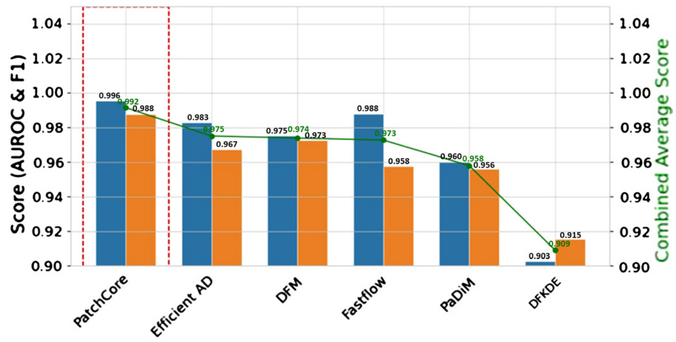


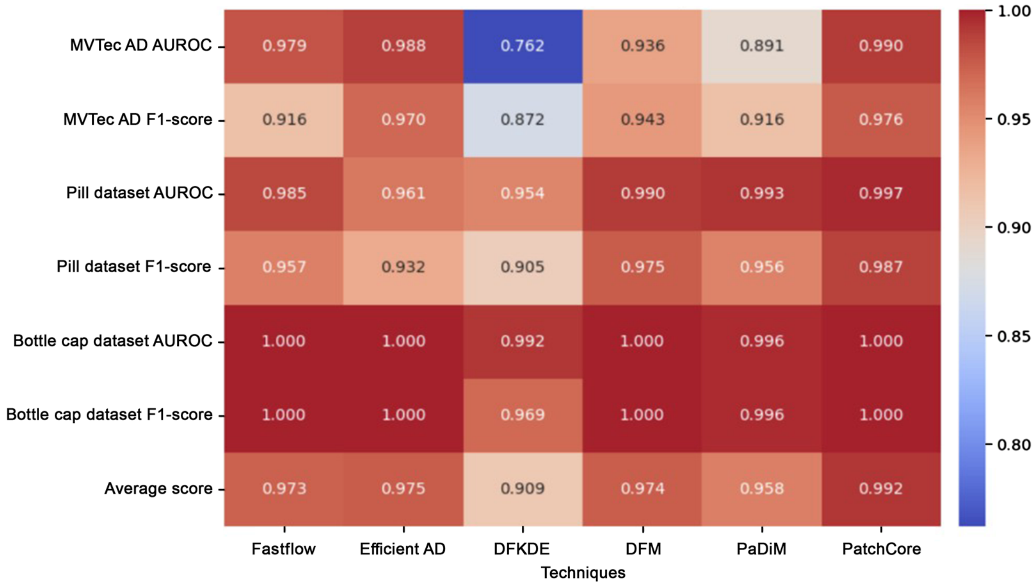
Figure 13
Average performance comparison



5. Discussion

Integrating PatchCore into IAD systems represents a significant advancement in manufacturing quality control, yet it also presents several challenges. In this work, combining CNN with patch-based feature extraction has notably improved the identification and localization of anomalies across high-resolution images from the MVTec AD dataset. The methodology shows that training solely on defect-free images and then evaluating images with anomalies can achieve high detection accuracy while effectively reducing false positives compared to traditional methods. This represents a leap from manual and conventional automated visual inspections, where human error and rigid rule-based systems often fall short in adapting to subtle or variable defect patterns. However, deploying PatchCore in real-world industrial settings does present integration challenges. The necessity for extensive preprocessing of high-resolution images and the reliance on high-performance computational facilities for initial training may hinder immediate onsite applicability in legacy production lines. Additionally, ensuring balanced datasets to prevent model overfitting remains a critical challenge, as imbalances in defect representation can bias the detection performance. The patch-based strategy excels in efficiently localizing anomalies, but its dependence on optimal patch size and configuration demands further research to standardize these parameters across different manufacturing contexts. Future improvements involve incorporating adaptive augmentation techniques and leveraging ensemble methods to strengthen the robustness of AD. Despite these challenges, the potential for increased operational efficiency, cost savings, and enhanced product quality positions PatchCore as a transformative tool for industrial applications. This work demonstrates significant technical advancements in defect detection and lays the groundwork for future studies to integrate AI-driven solutions more seamlessly into existing industrial quality

Figure 14
AD performance heatmap



control frameworks, ultimately fostering a new era of intelligent manufacturing practices.

The experimental results show that the standard implementation of PatchCore has the state-of-the-art performance on the selected benchmarks. In particular, the selection of the coreset subsampling algorithm and configuration of the memory bank (patch size, coreset sampling, and stride) have a direct influence on the computational efficiency/memory consumption/detection accuracy trade-off. Future work will include a more rigorous ablation analysis of the effect of these hyperparameters, the use of sampling algorithms other than the standard greedy coreset approach, and the optimization of the memory bank construction for specific defect types and modalities of image data. This will not only make the model effective but also in order to be fine-tuned for practical deployment scenarios.

One of the challenges identified in this research is the detection of small and subtle anomalies, which still present a problem for automated systems and cannot be identified without a high false-positive ratio. While patch-based correlation of PatchCore provides a certain level of granularity that favorably affects small defect detection, the standard implementation used here did not feature specific architectural changes (e.g., multi-scale feature extraction or dedicated attention mechanisms) designed specifically to improve performance on such fine-grained defects. To explicitly address this limitation, future work will address this limitation by incorporating multi-scale analysis pipelines and hierarchical memory banks and validate them on specialized benchmarks containing a higher proportion of minute anomalies to quantitatively show improvements.

The evaluation is that it is based on a single, dataset-wide threshold, which is chosen as a result of F1-score maximization on the validation set. A false negative (missed defect) could lead to compromising the product safety and reputation of the brand, whereas a false positive (false alarm) leads to unnecessary downtime and the cost of inspection. More robust thresholding strategies that are application-aware must therefore be a future priority. This involves using precision-recall curves to choose a threshold that satisfies certain precision (reduce false alarms) or recall (catch all defects) criteria, or using cost-sensitive learning to explicitly reduce financial loss, or even adaptive thresholding techniques that can adjust dynamically to changing data distributions on the production line.

5.1. Interpretability and anomaly justification

Integration of Gradient-weighted Class Activation Mapping (Grad-CAM): While PatchCore determines which anomalies are identified by the distance of features from the patch, Grad-CAM can be used to determine which features within the pretrained (WideResNet-50-2) backbone are most important for the decision.

The process would have worked as follows:

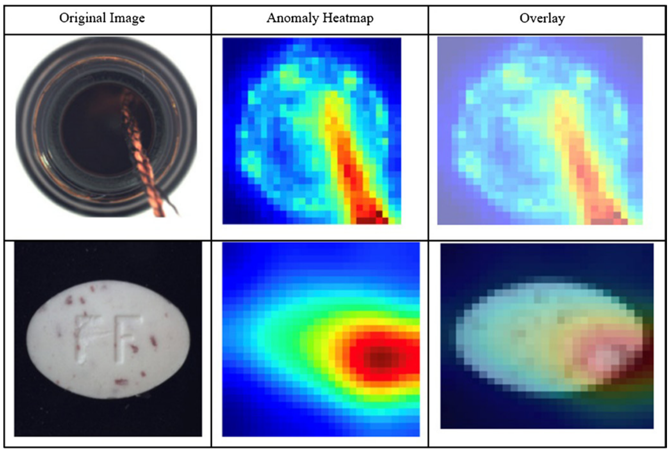
Generate PatchCore anomaly map: This gives the first localization of anomalous areas.

Calculate Grad-CAM: For the best anomalous image patch, calculate the gradients of the anomaly score (distance to the nearest neighbor in the memory bank) with regard to the feature maps from the last convolution layer in the backbone CNN.

Fused visualization: The resulting Grad-CAM heatmap is then overlaid on the original image, showing certain textures and structures the model thought to be the most abnormal, as presented in Figure 15.

On the left is the original image of the object, showing a close-up of the bottle and pill. In the center is the anomaly heatmap, which highlights the areas of the image with detected anomalies. These anomalies are represented in a color gradient indicating higher levels of detected irregularity. On the right, the heatmap is overlaid onto the original image to visually emphasize where the anomalies are located, helping to identify defects or irregularities in the object.

Figure 15
Bottle and pill anomaly heatmap and overlay



5.2. Challenges and solutions in IAD

Vision-based AD is being integrated into various domains to address inspection and quality control issues. However, automatic detection brings new obstacles. These problems can be classified as scientific or industrial. Scientific domains focus on the detection model and its operation, whereas industrial issues focus on the practical use of automated detection in industry. Key challenges include real-time detection, sophisticated inspections, annotations, insufficient datasets, data management, and system integration.

1) **Real-time inspection and decision-making:** Real-time model needs adaptability, minimal latency, and seamless processing. It is crucial for detecting systems to keep up with current processes. Due to the dynamic nature of industrial processes, new anomaly samples may arise during real-world operations. Due to their fixed dataset training, offline learning methods may not accurately recognize or categorize new occurrences. To address these issues, consider adding edge processing. Unlike cloud computing, edge processing reduces latency and speeds up decision-making by bringing data closer to the inspection system. Lightweight defect detection algorithms like Bergmann et al.'s [60] also balance speed and accuracy while reducing computing parameters. However, there is potential for developing speedier inspection and more efficient, lightweight fault detection models.
2) **Small defects and annotation:** When using automated inspection, addressing small problems that are not visible to the naked eye is challenging. Regardless of size, small faults can compromise a structure's integrity and cause failure. Detecting minor flaws is challenging due to their size in the broader region of interest. To mitigate this issue, inspection models should have a low false-positive rate [62], reducing the number of defective occurrences misclassified as non-defective. Annotated datasets and minor flaws pose a continuing problem [63]. However, annotation of fine detail across several images is expensive and time-consuming. Approaches using generative models aim to overcome these issues. These techniques use generative networks to synthesize big labeled datasets using a minimal amount of manually labeled data [64].

3) **Unbalancing and insufficient datasets:** Obtaining sufficient industrial datasets is a significant difficulty for learning-based methods. A bias toward plentiful classes emerges from an imbalance between faulty and non-defective samples, with defective examples being harder to obtain. This may impact performance metrics, as high detection scores may be more typical of the overall class and not correspond with minority samples. Additionally, insufficient training datasets reduce detection accuracy as the model lacks experience in all categories. Small and imbalanced datasets can be addressed by increasing sample size using data augmentation, generative networks, and synthetic approaches [65]. Few-shot learning is a prominent method for addressing data scarcity by training models on a small number of samples. However, dataset reliance remains a key feature of IAD and is constantly being enhanced.
4) **Data quality and management:** This branch also manages vast amounts of data. Data samples must be properly collected, stored, and distributed for effective solutions. The inspection system uses feedback from evaluations, user inputs, and neighboring systems. Quality data is essential for inspection, as low-quality data can inhibit it [66, 61, 67]. To overcome quality issues, incorporate appropriate sensing and illumination components. Alternative methods include training, which enhances model resilience and performance under diverse lighting and imaging situations. Additionally, development toward Industry 5.0 will reduce technical and logistical data-collecting obstacles [68].

6. Conclusion and Future Scope

This study demonstrates the transformative potential of the PatchCore algorithm in IAD by leveraging high-resolution images from the MVTec AD dataset. Investigation confirms that integrating CNN with patch-based feature extraction significantly enhances defect detection capabilities for various anomalies, including scratches, dents, and contaminations. The presented method not only improves detection accuracy and reduces false positives but also streamlines the quality inspection process, thereby offering substantial benefits in efficiency and cost-effectiveness over traditional manual and automated inspection systems. Despite the evident advantages, our work reveals critical challenges, such as the demands for extensive computational resources during the training phase and difficulties in integrating advanced AI systems with legacy manufacturing infrastructures. These limitations underscore the need for further research into optimizing processing pipelines, calibrating detection thresholds, and refining model deployment strategies for diverse industrial environments. The study represents a pivotal step toward infusing smart technology into quality control practices, setting the stage for resilient, future-ready manufacturing systems.

IAD using PatchCore presents a clear path toward transformative advancements in manufacturing quality control. Incorporating transfer learning and unsupervised learning techniques to enhance detection capabilities further can help to address challenges such as limited defect-labeled data and heterogeneous production environments. Future efforts will focus on integrating this model with Internet of Things (IoT) technology, enabling real-time data acquisition from sensors and edge devices for immediate anomaly identification and response. Additionally, exploring multimodal data fusion—combining visual, acoustic, and thermal imaging—may yield improved accuracy in detecting subtle defects under varying operational conditions. Research can also investigate the optimization of lightweight, energy-efficient network architectures to ensure that robust AD algorithms are deployable on resource-constrained devices. Emphasis on adaptive learning methods to continually update detection models in response to changing manufacturing processes promises more resilient and scalable systems. Moreover, collaborations with industry stakeholders will be vital in tailoring these AI solutions to diverse production lines, paving the way for fully integrated, smart quality control systems. These efforts will catalyze the evolution of sustainable, intelligent manufacturing practices that set new efficiency and product reliability benchmarks. These efforts promise to establish robust, scalable AI solutions for future smart manufacturing.

The primary objective of our paper is to improve the accuracy on the public benchmarks (MVTec AD, pill, bottle cap). Despite the vital role that public datasets play in establishing and benchmarking current algorithms, they are notoriously lacking the sharp class imbalances, noise characteristics, and defect modes that are prevalent in manufacturing datasets that are currently present in the real world. For this, the preprocessing steps included resizing the image, normalization, and center cropping of the image. In order to achieve this performance, further implementations are carried out on choosing the patch size, coreset sampling ratio, and layers 2 and 3. The above features make the usage of the existing benchmark dataset more novel to test the performance.

The further research will therefore be oriented toward quantifying the model's inference latency and throughput on industry-standard edge computing devices (e.g., NVIDIA Jetson AGX Orin) to validate the suitability of the model for real-time inspection. It also includes pilot integration through working with industry partners to implement the system on a live production line for a limited pilot study. This will permit the collection of the most important real-life measures: false-positive rate (FPR), which has a direct impact on production downtime, and the false-negative rate (FNR), which has an impact on quality control.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available at https://www.mvtec.com/company/research/datasets/mvtec-ad.

Author Contribution Statement

Shalini Kumari: Conceptualization, Software, analysis, Resources, Data curation, Writing – original draft. **Chander Prabha:** Methodology, Validation, Investigation, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Shan, B., Zheng, S., & Ou, J. (2016). A stereovision-based crack width detection approach for concrete surface assessment. *KSCE Journal of Civil Engineering*, *20*(2), 803–812. https://doi.org/10.1007/s12205-015-0461-6

[2] Fujita, Y., & Hamamoto, Y. (2011). A robust automatic crack detection method from noisy concrete surfaces. *Machine Vision and Applications*, *22*(2), 245–254. https://doi.org/10.1007/s00138-009-0244-5

[3] Hyun, J., Kim, S., Jeon, G., Kim, S. H., Bae, K., & Kang, B. J. (2024). ReConPatch: Contrastive patch representation learning for industrial anomaly detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2041–2050. https://doi.org/10.1109/WACV57701.2024.00205

[4] Peng, T., Zheng, Y., Zhao, L., & Zheng, E. (2024). Industrial product surface anomaly detection with realistic synthetic anomalies based on defect map prediction. *Sensors*, *24*(1), 264. https://doi.org/10.3390/s24010264

[5] Jiang, Z., Zhang, Y., Wang, Y., Li, J., & Gao, X. (2024). FR-PatchCore: An industrial anomaly detection method for improving generalization. *Sensors*, *24*(5), 1368. https://doi.org/10.3390/s24051368

[6] Wen, P., Gao, X., Wang, Y., Li, J., & Luo, L. (2023). Normalizing flow-based industrial complex background anomaly detection. *Journal of Sensors*, *2023*(1), 6690190. https://doi.org/10.1155/2023/6690190

[7] Liu, Z., Zhou, Y., Xu, Y., & Wang, Z. (2023). SimpleNet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20402–20411. https://doi.org/10.1109/CVPR52729.2023.01954

[8] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., & Gehler, P. (2022). Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14298–14308. https://doi.org/10.1109/CVPR52688.2022.01392

[9] Ishida, K., Takena, Y., Nota, Y., Mochizuki, R., Matsumura, I., & Ohashi, G. (2023). SA-PatchCore: Anomaly detection in dataset with co-occurrence relationships using self-attention. *IEEE Access*, *11*, 3232–3240. https://doi.org/10.1109/ACCESS.2023.3234745

[10] Tang, Q., & Jung, H. (2023). Reliable anomaly detection and localization system: Implications on manufacturing industry. *IEEE Access*, *11*, 114613–114622. https://doi.org/10.1109/ACCESS.2023.3324314

[11] Choi, B., & Jeong, J. (2022). ViV-Ano: Anomaly detection and localization combining vision transformer and variational autoencoder in the manufacturing process. *Electronics*, *11*(15), 2306. https://doi.org/10.3390/electronics11152306

[12] Bozcan, I., Korndorfer, C., Madsen, M. W., & Kayacan, E. (2022). Score-based anomaly detection for smart manufacturing systems. *IEEE/ASME Transactions on Mechatronics*, *27*(6), 5233–5242. https://doi.org/10.1109/TMECH.2022.3169724

[13] Jezek, S., Jonak, M., Burget, R., Dvorak, P., & Skotak, M. (2021). Deep learning-based defect detection of metal parts: Evaluating current methods in complex conditions. In *2021 13th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, 66–71. https://doi.org/10.1109/ICUMT54235.2021.9631567

[14] Defard, T., Setkov, A., Loesch, A., & Audigier, R. (2021). PaDiM: A patch distribution modelling framework for anomaly detection and localization. In *25th International Conference on Pattern Recognition Workshops and Challenges*, 475–489. https://doi.org/10.1007/978-3-030-68799-1_35

[15] Zhao, H., Jia, J., & Koltun, V. (2020). Exploring Self-attention for Image Recognition. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10073–10082. https://doi.org/10.1109/CVPR42600.2020.01009

[16] Kumari, S., Prabha, C., Karim, A., Hassan, Md. M., & Azam, S. (2024). A comprehensive investigation of anomaly detection methods in deep learning and machine learning: 2019–2023. *IET Information Security*, *2024*(1), 8821891. https://doi.org/10.1049/2024/8821891

[17] Kumari, S., & Prabha, C. (2024). A comprehensive review of deep anomaly detection techniques- An analysis. In *2024 IEEE 9th International Conference for Convergence in Technology*, 1–6. https://doi.org/10.1109/I2CT61223.2024.10543335

[18] Lin, Y., Chang, Y., Tong, X., Yu, J., Liotta, A., Huang, G.,…, & Zhang, W. (2025). A survey on RGB, 3D, and multimodal approaches for unsupervised industrial image anomaly detection. *Information Fusion*, *121*(C), 103139. https://doi.org/10.1016/j.inffus.2025.103139

[19] Liang, Y., Hu, Z., Huang, J., Di, D., Su, A., & Fan, L. (2025). ToCoAD: Two-stage contrastive learning for industrial anomaly detection. *IEEE Transactions on Instrumentation and Measurement*, *74*, 1–9. https://doi.org/10.1109/TIM.2025.3545987

[20] Zonnenshain, A., & Kenett, R. S. (2020). Quality 4.0—The challenging future of quality engineering. *Quality Engineering*, *32*(4), 614–626. https://doi.org/10.1080/08982112.2019.1706744

[21] Ciberlin, J., Grbic, R., Teslic, N., & Pilipovic, M. (2019). Object detection and object tracking in front of the vehicle using front view camera. In *2019 Zooming Innovation in Consumer Technologies Conference*, 27–32. https://doi.org/10.1109/ZINC.2019.8769367

[22] Bansal, M., Kumar, M., & Kumar, M. (2021). 2D object recognition: A comparative analysis of SIFT, SURF and ORB feature descriptors. *Multimedia Tools and Applications*, *80*(12), 18839–18857. https://doi.org/10.1007/s11042-021-10646-0

[23] Zhou, W., Gao, S., Zhang, L., & Lou, X. (2020). Histogram of oriented gradients feature extraction from raw Bayer pattern images. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *67*(5), 946–950. https://doi.org/10.1109/TCSII.2020.2980557

[24] Dhakshina Kumar, S., Esakkirajan, S., Bama, S., & Keerthiveena, B. (2020). A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier. *Microprocessors and Microsystems*, *76*, 103090. https://doi.org/10.1016/j.micpro.2020.103090

[25] Ahmed, A., Jalal, A., & Kim, K. (2019). Region and decision tree-based segmentations for multi-objects detection and classification in outdoor scenes. In *2019 International Conference on Frontiers of Information Technology*, 209–2095. https://doi.org/10.1109/FIT47737.2019.00047

[26] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S.,…& Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics*, *10*(20), 2470. https://doi.org/10.3390/electronics10202470

[27] Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A.,…, & Socher, R. (2021). Deep learning-enabled medical computer vision. *npj Digital Medicine*, *4*(1), 5. https://doi.org/10.1038/s41746-020-00376-2

[28] Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I.,…, & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, *35*(5), 1285–1298. https://doi.org/10.1109/TMI.2016.2528162

[29] Xie, G., Wang, J., Liu, J., Lyu, J., Liu, Y., Wang, C.,…, & Jin, Y. (2024). IM-IAD: Industrial image anomaly detection benchmark in manufacturing. *IEEE Transactions on Cybernetics*, *54*(5), 2720–2733. https://doi.org/10.1109/TCYB.2024.3357213

[30] Bertolini, M., Mezzogori, D., Neroni, M., & Zammori, F. (2021). Machine Learning for industrial applications: A comprehensive

literature review. *Expert Systems with Applications*, *175*, 114820. https://doi.org/10.1016/j.eswa.2021.114820

[31] Ahmad, H. M., & Rahimi, A. (2022). Deep learning methods for object detection in smart manufacturing: A survey. *Journal of Manufacturing Systems*, *64*, 181–196. https://doi.org/10.1016/j.jmsy.2022.06.011

[32] Yasuda, Y. D. V., Cappabianco, F. A. M., Martins, L. E. G., & Gripp, J. A. B. (2022). Aircraft visual inspection: A systematic literature review. *Computers in Industry*, *141*, 103695. https://doi.org/10.1016/j.compind.2022.103695

[33] Konstantinidis, F. K., Myrillas, N., Tsintotas, K. A., Mouroutsos, S. G., & Gasteratos, A. (2023). A technology maturity assessment framework for Industry 5.0 machine vision systems based on systematic literature review in automotive manufacturing. *International Journal of Production Research*. Advance online publication. https://doi.org/10.1080/00207543.2023.227058

[34] Jha, S. B., & Babiceanu, R. F. (2023). Deep CNN-based visual defect detection: Survey of current literature. *Computers in Industry*, *148*, 103911. https://doi.org/10.1016/j.compind.2023.103911

[35] Kumari, S., & Prabha, C. (2023). A comprehensive review on anomaly detection in images: Challenges and future research directions. In *2023 IEEE North Karnataka Subsection Flagship International Conference*, 1–6. https://doi.org/10.1109/NKCon59507.2023.10396507

[36] Ameri, R., Hsu, C.-C., & Band, S. S. (2024). A systematic review of deep learning approaches for surface defect detection in industrial applications. *Engineering Applications of Artificial Intelligence*, *130*, 107717. https://doi.org/10.1016/j.engappai.2023.107717

[37] Hojjati, H., Ho, T. K. K., & Armanfard, N. (2024). Self-supervised anomaly detection in computer vision and beyond: A survey and outlook. *Neural Networks*, *172*, 106106. https://doi.org/10.1016/j.neunet.2024.106106

[38] Trilles, S., Hammad, S. S., & Iskandaryan, D. (2024). Anomaly detection based on artificial intelligence of things: A systematic literature mapping. *Internet of Things*, *25*, 101063. https://doi.org/10.1016/j.iot.2024.101063

[39] Shukla, V., Shukla, A., S. K., S. P., & Shukla, S. (2025). A systematic survey: Role of deep learning-based image anomaly detection in industrial inspection contexts. *Frontiers in Robotics and AI*, *12*, 1554196. https://doi.org/10.3389/frobt.2025.1554196

[40] Kraitem, Z. (2025). A comparative study of text-to-image synthesis techniques using generative adversarial networks. *American Journal of Neural Networks and Applications*, *11*(1), 24–30. https://doi.org/10.11648/j.ajae.20231001.13

[41] Bucci, S., D'Innocente, A., Liao, Y., Carlucci, F. M., Caputo, B., & Tommasi, T. (2022). Self-supervised learning across domains. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(9), 5516–5528. https://doi.org/10.1109/TPAMI.2021.3070791

[42] Hendrycks, D., Mazeika, M., Kadavath, S., & Song, D. (2019). Using self-supervised learning can improve model robustness and uncertainty. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 15663–15674.

[43] Adam, A., Rivlin, E., Shimshoni, I., & Reinitz, D. (2008). Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(3), 555–560. https://doi.org/10.1109/TPAMI.2007.70825

[44] Guo, J., Lu, S., Jia, L., Zhang, W., & Li, H. (2023). ReContrast: Domain-specific anomaly detection via contrastive reconstruction. In *Proceedings of the 27th Conference on Neural Information Processing Systems*, 10721–10740.

[45] Tang, Q., Cao, G., & Jo, K.-H. (2021). Integrated feature pyramid network with feature aggregation for traffic sign detection. *IEEE Access*, 9, 117784–117794. https://doi.org/10.1109/ACCESS.2021.3106350

[46] Liang, Y., Zhang, J., Zhao, S., Wu, R., Liu, Y., & Pan, S. (2023). Omni-frequency channel-selection representations for unsupervised anomaly detection. *IEEE Transactions on Image Processing*, *32*, 4327–4340. https://doi.org/10.1109/TIP.2023.3293772

[47] Rudolph, M., Wandt, B., & Rosenhahn, B. (2021). Same same but differnet: Semi-supervised defect detection with normalizing flows. In *2021 IEEE Winter Conference on Applications of Computer Vision*, 1906–1915. https://doi.org/10.1109/WACV48630.2021.00195

[48] Venkataramanan, S., Peng, K.-C., Singh, R. V., & Mahalanobis, A. (2020). Attention guided anomaly localization in images. In *ECCV 2020: 16th European Conference on Computer Vision*, 485–503. https://doi.org/10.1007/978-3-030-58520-4_29

[49] Kong, S., & Ramanan, D. (2025). OpenGAN: Open-set recognition via open data generation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *47*(5), 3233–3243. https://doi.org/10.1109/TPAMI.2022.3184052

[50] Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2020). Uninformed students: Student-teacher anomaly detection with discriminative latent embeddings. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4182–4191. https://doi.org/10.1109/CVPR42600.2020.00424

[51] Wang, N., Zhang, Z., Zhao, X., Miao, Q., Ji, R., & Gao, Y. (2019). Exploring high-order correlations for industry anomaly detection. *IEEE Transactions on Industrial Electronics*, *66*(12), 9682–9691. https://doi.org/10.1109/TIE.2019.2907441

[52] Sabokrou, M., Khalooei, M., Fathy, M., & Adeli, E. (2018). Adversarially learned one-class classifier for novelty detection. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3379–3388. https://doi.org/10.1109/CVPR.2018.00356

[53] Wang, Q., Yuan, Z., Du, Q., & Li, X. (2019). GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 3–13. https://doi.org/10.1109/TGRS.2018.2849692

[54] Yao, R., Liu, C., Zhang, L., & Peng, P. (2019). Unsupervised anomaly detection using Variational Auto-Encoder based feature extraction. In *2019 IEEE International Conference on Prognostics and Health Management*, 1–7. https://doi.org/10.1109/ICPHM.2019.8819434

[55] Staar, B., Lütjen, M., & Freitag, M. (2019). Anomaly detection with convolutional neural networks for industrial surface inspection. *Procedia CIRP*, *79*, 484–489. https://doi.org/10.1016/j.procir.2019.02.123ote.12394

[56] Jing, J., Ma, H., & Zhang, H. (2019). Automatic fabric defect detection using a deep convolutional neural network. *Coloration Technology*, *135*(3), 213–223. https://doi.org/10.1111/cote.12394

[57] Perera, P., & Patel, V. M. (2019). Learning deep features for one-class classification. *IEEE Transactions on Image Processing*, *28*(11), 5450–5463. https://doi.org/10.1109/TIP.2019.2917862

[58] Ullah, W., Khan, S. U., Kim, M. J., Hussain, A., Munsif, M., Lee, M. Y.,…, & Baik, S. W. (2024). Industrial defective chips detection using deep convolutional neural network with inverse feature matching mechanism. *Journal of Computational Design and Engineering*, *11*(3), 326–336. https://doi.org/10.1093/jcde/qwae019

[59] Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D., & Steger, C. (2021). The MVTec anomaly detection dataset: A comprehensive real-world dataset for unsupervised anomaly detection. *International Journal of Computer Vision*, *129*(4), 1038–1059. https://doi.org/10.1007/s11263-020-01400-4

[60] Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD—A comprehensive real-world dataset for unsupervised anomaly detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9584–9592. https://doi.org/10.1109/CVPR.2019.00982

[61] Li, D., Ling, H., Kim, S. W., Kreis, K., Fidler, S., & Torralba, A. (2022). BigDatasetGAN: Synthesizing imagenet with pixel-wise annotations. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 21298–21308. https://doi.org/10.1109/CVPR52688.2022.02064

[62] Akcay, S., Ameln, D., Vaidya, A., Lakshmanan, B., Ahuja, N., & Genc, U. (2022). Anomalib: A deep learning library for anomaly detection. In *2022 IEEE International Conference on Image Processing*, 1706–1710. https://doi.org/10.1109/ICIP46576.2022.9897283

[63] Malesa, M., & Rajkiewicz, P. (2021). Quality control of pet bottles caps with dedicated image calibration and deep neural networks. *Sensors*, *21*(2), 501. https://doi.org/10.3390/s21020501

[64] Yuan, L., Chen, Y., Tang, H., Gao, R., & Wu, W. (2024). A lightweight deep-learning algorithm for welding defect detection in new energy vehicle battery current collectors. *IEEE Sensors Journal*, *24*(13), 21655–21668. https://doi.org/10.1109/JSEN.2024.3398769

[65] Dong, X., Taylor, C. J., & Cootes, T. F. (2019). Small defect detection using convolutional neural network features and random forests. In *ECCV 2018 Workshops: European Conference on Computer Vision Workshops*, 398–412. https://doi.org/10.1007/978-3-030-11018-5_35

[66] Mei, S., Cheng, J., He, X., Hu, H., & Wen, G. (2022). A novel weakly supervised ensemble learning framework for automated pixel-wise industry anomaly detection. *IEEE Sensors Journal*, *22*(2), 1560–1570. https://doi.org/10.1109/JSEN.2021.3131908

[67] Liu, S., Ni, H., Li, C., Zou, Y., & Luo, Y. (2024). DefectGAN: Synthetic data generation for EMU defects detection with limited data. *IEEE Sensors Journal*, *24*(11), 17638–17652. https://doi.org/10.1109/JSEN.2024.3386711

[68] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., & Wang, C. (2023). Multimodal industrial anomaly detection via hybrid fusion. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8032–8041. https://doi.org/10.1109/CVPR52729.2023.00776

How to Cite: Kumari, S., & Prabha, C. (2026). Anomaly Detection Utilizing PatchCore for Reimagining Industrial Visual Inspection. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA52026321