A Deep Dive into Using CNNS for Spotting Anomalies in Industrial Visual Checks: Methods and Real-World Applications Explored

ADITYA KINNORI[1], ANIKET TRIPATHI[2]
[1, 2]*Nirma University, India*

Abstract- Recent advances in deep learning have made it possible for production lines to adopt automated and accurate anomaly detection in industrial visual inspections. Convolutional Neural Networks (CNNs), in particular, have shown superior performance over other models due to their ability to capture structured patterns in visual data. This paper presents a detailed survey of CNN-based approaches for identifying anomalies in industrial settings. Techniques are grouped into supervised, unsupervised, and self-supervised categories, with a focus on their strengths, limitations, and common use cases. The review also covers hybrid approaches that combine CNNs with generative models such as autoencoders and GANs to improve performance in data-scarce environments. A thorough catalog of available datasets is included, along with evaluation methods and comparative results across different CNN models in real-world industrial scenarios. Key deployment challenges are discussed, including limited data availability, domain shifts, model interpretability, and the need for real-time processing. Additionally, the paper highlights emerging trends and recommends future directions such as integrating Vision Transformers, leveraging contrastive learning, and prioritizing edge deployment. This survey aims to support professionals involved in building, implementing, or refining CNN-based anomaly detection systems in modern industrial operations.

Index Terms- CNN, anomaly detection, industrial visual inspection, deep learning, autoencoder, GAN, real-time inspection, defect detection.

I. INTRODUCTION

Automation has become a cornerstone of quality control in modern manufacturing, driven by the need for high precision, consistency, and reduced dependency on human labor. Traditional visual inspections, which rely heavily on human operators, often struggle with issues like fatigue, subjective judgment, and difficulty scaling to large production volumes. In response, computer vision systems powered by Convolutional Neural Networks (CNNs) are increasingly being adopted due to their proven ability to recognize complex visual patterns across varied data formats.

CNNs have achieved remarkable results in tasks such as image classification, object detection, and segmentation, which has led to their application in industrial anomaly detection for identifying surface defects, inspecting printed circuit boards, and analyzing semiconductor wafers (Christiansen et al., 2016; Chen et al., 2018). Their strength lies in detecting both spatial features and textures, enabling them to identify subtle product deviations that may indicate flaws.

However, spotting anomalies in industrial contexts remains challenging. The unpredictable nature and rarity of defects often result in highly imbalanced datasets, which hampers the performance of supervised learning methods (Sabokrou et al., 2017). This has shifted research interest toward unsupervised and semi-supervised approaches, where CNNs are trained only on normal samples to model standard behavior. Anomalies are then inferred from reconstruction errors or discrepancies in learned representations (Larsen et al., 2017; Kim et al., 2017).

While earlier machine learning models like SVMs and decision trees once played a major role, their reliance on handcrafted features limited their effectiveness for analyzing complex imagery (Assendorp & Deep, 2017). The emergence of deep learning has allowed for end-to-end learning from raw image data, significantly improving performance in unstructured visual domains.

Nonetheless, implementing CNN-based anomaly detection in real-world industrial environments remains difficult. Key obstacles include high computational demands, the need for real-time inference, adaptation across diverse production lines, and the general lack of interpretability in deep models. Tackling these challenges is critical for developing robust and scalable inspection systems.

This paper offers an in-depth review of CNN-based methods for industrial visual inspection. It classifies current techniques into supervised, unsupervised, and self-supervised categories, examines widely used benchmark datasets, and outlines common evaluation strategies. The review also identifies unresolved issues and proposes potential research directions to advance the field further.

II. BACKGROUND AND THEORETICAL FOUNDATIONS

CNNs are able to help with computer vision tasks like image classification and anomaly detection, thanks to their strong power to learn different levels of importance in image features. Because faults are usually not easy to spot and are small and limited in industrial visual inspection, this way of seeing is very important. This part sets up a base for learning how CNNs help detect anomalies, clarifies the types of approaches used and discusses what features of a CNN are important for identifying anomalies in challenging industrial situations.

III. CONVOLUTIONAL NEURAL NETWORKS IN VISUAL REPRESENTATION LEARNING

These networks take their inspiration from our brains' visual thinking, using layers of filters to pick out useful details from each image. In the early 1990s, CNNs started to gain recognition, but it was AlexNet in 2012 that led to their greater popularity following its good results in the ImageNet challenge (Yamashita et al., 2018). Convolutional, activation, pooling and fully connected layers are generally part of a CNN architecture. The special design of CNNs allows complex patterns in images to be handled with very little effort by humans in deciding which parts of the data to use. CNNs in industrial anomaly detection are capable of telling apart the textures of flaws and those of undamaged surfaces. Instead of using handmade features, CNNs adapt their weights on their own to pay attention only to essential and constant aspects (Sabokrou et al., 2017).

IV. ANOMALY DETECTION: DEFINITIONS AND STRATEGIES

Anomaly detection is about finding instances in data that are very different from the standard amount. To do this, inspectors check for scratches, cracks and wrong alignments, as well as contamination on the outside of the product. Because these defects come up occasionally and vary in their appearance, they are hard to catch with simple inspectors that depend on rules. In their work, Sabokrou et al. (2017) divide anomaly detection into three models: supervised, semi-supervised and unsupervised. When using supervised learning, both correct and faulty samples are used, but semi-supervised and unsupervised approaches mostly depend on normal data for training.

Anomaly detectors with supervised CNNs are trained using images with proper classifications. Still, because there aren't many defective samples, classifying the data becomes a huge challenge for the classifiers. Instead, when using these methods, only normal noise values are modeled and anomalies are spotted while inferring new data. Because it is difficult to collect many diverse defect samples in industry, this becomes very useful in that area (Larsen et al., 2016; Kim et al., 2017).

V. CNN ARCHITECTURES IN ANOMALY DETECTION SYSTEMS

Different CNN designs are used for different functions in detecting anomalies. The use of deep hierarchical layers in VGGNet and ResNet by classical architectures is valuable for identifying even the most varied defects (Assendorp & Deep, 2017). Autoencoders in unsupervised situations feed image data through CNN encoders which convert the data into latent features. The decoder restores the original image from the latent features. Greater reconstruction error may mean there is an abnormality present. VAEs and GANs also work on this idea by acquiring probabilistic and adversarial distributions, leading to better findings of small irregularities in detailed industrial pictures (Larsen et al., 2016; Kim et al., 2017).

Architecture	Year	Type	Typical Use in Anomaly Detection	Strengths
LeNet-5	1998	Supervised	Basic image classification	Lightweight, effective on small images
AlexNet	2012	Supervised	Feature extraction for defect classification	Deep filters, strong generalization
VGG-16/VGG-19	2014	Supervised	Fine-grained defect classification	High accuracy, simple architecture
ResNet (e.g., ResNet-50)	2016	Supervised/Unsupervised	Deep residual learning for complex textures	Solves vanishing gradient problem
CNN Autoencoder	2011	Unsupervised	Image reconstruction, anomaly scoring	Useful when no defect labels are available
GAN	2014	Unsupervised	Adversarial modeling, synthetic defect detection	Captures distribution shifts effectively

Table 1. Comparative roles and strengths of common CNN architectures.

VI. CNN-BASED FEATURE LEARNING VS. TRADITIONAL METHODS

Classical computer vision saw the frequent use of Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG) and Gabor filters, but that changed when CNNs became popular (Assendorp & Deep, 2017). While the approaches helped in clean tests, they did not work on surfaces covered with different textures used by industry. Instead, CNNs can discover the main qualities needed for finding defects which means the systems can be used more widely and effectively. Figure 1 makes it clear how a regular feature-based approach for inspection works differently from a CNN-based visual anomaly detection process.

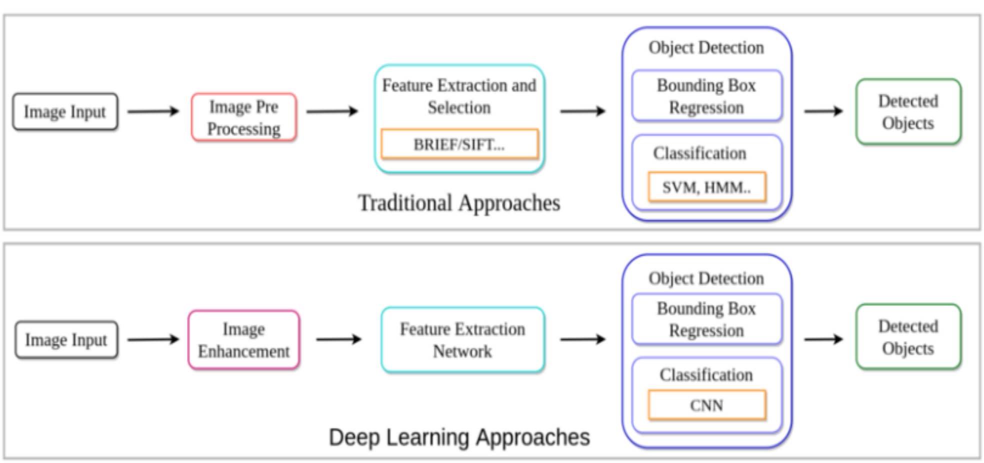

Figure 1. Traditional vs. CNN-Based Anomaly Detection in Visual Inspection

VII. CHALLENGES IN CNN-BASED VISUAL INSPECTION

Nevertheless, CNNs cannot overcome all of their disadvantages. Training a deep CNN well is often hard because many samples are needed, but flaws only happen rarely. In some cases, if the network picks up just the random differences of the training data, it will be easily fooled and won't do well in general. Also, how CNNs work makes it very hard to know why they predict what they do—an important worry regarding critical industrial problems. Although Grad-CAM and saliency maps can increase a model's transparency, their usefulness in everyday problems is still being studied (Sabokrou et al., 2017).

VIII. METHODOLOGICAL TAXONOMY

CNNs have made anomaly detection in industrial visual inspection a broad research interest with a variety of approaches. All approaches rest on different beliefs about how available the data is, how often anomalies arrive and what the system is set to do. We now provide a taxonomy of CNN-based approaches to detecting anomalies, organizing them into supervised, unsupervised, self-supervised and hybrid varieties. To understand the balance between detection, data load and deploying these methods in factories, this classification is important.

IX. SUPERVISED CNN-BASED METHODS

The first CNN uses in anomaly detection came as supervised algorithms and typically worked with sets of images both with and without defects. To use these approaches, CNNs are built as typical classifiers using cross-entropy or comparable loss functions, labeling each image with a left-out class or a binary normal/defect label. Many users pick AlexNet, VGGNet and ResNet because they work very smoothly during image classification (Yamashita et al., 2018). The major challenge for supervised approaches in industry is that the amount of normal data is dramatically higher than the number of defective samples. Not only are defective instances rare, but they can look quite different from one another which makes it challenging for CNNs to learn about all types of defects. Besides, bringing together and noting all the defective items needed for all types of goods can be very difficult and take a lot of time and money. Despite the fact that data augmentation improves accuracy in cases of class imbalance, it mostly fails in situations where objects within a class have a lot of differences (Assendorp & Deep, 2017).

X. SELF-SUPERVISED LEARNING APPROACHES

Because in industry most samples are good, unsupervised methods have been widely explored. Their training requires only normal samples, as they work to figure out the normal data pattern. Anomalies are found when there is variation from what is normal in the distribution. In common practice, the convolutional autoencoder takes an image, turns it into a latent code and reconstructs it to find anomalies by comparing the original and the reconstructed image (Madry et al., 2018). Dap variation autoencoders add probability modeling and generative adversarial networks use adversarial training to extend this idea. These models were found to perform strongly in detecting soft surface defects because they examine detailed variations in data (Larsen et al., 2016; Kim et al., 2017). Even so, issues with unstable training and assessing the quality of results still arise, mainly during detailed high-resolution scanning.

XI. HYBRID AND ONE-CLASS CNN MODELS

Supervised and unsupervised learning are contrasted by a third method, self-supervised learning. It finds ways to learn useful features by using tasks invented from unlabeled material, rather than relying on manually added labels. In industrial inspection, it has been shown that self-supervised CNNs can produce stable representations that highlight the structure and appearances of normal products (Sabokrou et al., 2017). When these embeddings are not followed, the data is marked as an anomaly. Thanks to these approaches, there's no need to see damaged material first to detect issues at a high level of accuracy.

Such models use more than one approach to learning to make anomaly detection more effective. One example is that CNNs used with one-class classification can create small representations of normal samples that can help spot strange or unusual patterns. Converting these images into probability distributions, convolutional encoders or adversarial learning have been explored to better handle shifts in distributions in visual inspection tasks (Kim et al., 2017; Sabokrou et al., 2017). They provide an effective middle ground between flexible unsupervised learning and the feature recognition in supervised CNNs. Such models are best used when making the difference between normal and unusual cases relies on the situation

Approach	Description	Key Models	Sample Use Cases	Strengths	Limitations
Supervised	Learns directly from labeled defect/non-defect data	AlexNet, VGG, ResNet	Scratch classification, PCB inspection	High accuracy with labeled data	Requires large annotated datasets
Unsupervised	Trained only on normal samples; detects outliers by deviation	Autoencoders, GANs, VAEs	Surface defect localization	No need for defect samples	Sensitive to noise and distribution shifts
Self-Supervised	Learns representations using pretext tasks	RotationNet, PuzzleNet	Fabric and texture anomaly detection	Label-efficient, robust features	Performance depends on pretext task
Hybrid/One-Class	Combines CNNs with OC-SVM or clustering loss	Deep SVDD, AE+GAN	Subtle defect classification	Effective for rare anomalies	Complex training and tuning

Table 2. Methodological Taxonomy of CNN-Based Anomaly Detection in Industrial Visual Inspection

XII. APPLICATIONS IN INDUSTRY

Using CNN-based anomaly detection in industrial visual inspection has made it easier to automate the process, improve accuracy when finding defects and lessen costs. Many industries including semiconductor manufacturing, auto assembly, textiles and drug packaging use CNNs to catch small, extreme and focused defects. They reveal how practical and adaptable CNNs are in many industries that require high accuracy, scale up and the ability to process quickly (Yamashita et al., 2018).

XIII. SEMICONDUCTOR AND ELECTRONICS MANUFACTURING

Computer vision experts in semiconductors and printed circuit boards started using CNNs for anomaly detection early, because the field requires strict quality and deals with challenging types of errors. Traditional approaches based on rules regularly skip over small problems such as missing part or cable breaks present in images under changing environment. Unlike kinematics, CNNs can master spatial patterns in clean or labeled data and therefore provide accurate and dependable detection. ResNet and Inception Deep CNNs have been applied with automated optical inspection (AOI) to make the systems more sensitive and efficient at real-time use. For quickly detecting and locating several types of defects, current applications make use of YOLOv5 and PyTorch architectures, both improved by GPU processing. Almost all systems include an AOI camera, a way to save captured images and a CNN module that permits real-time inference.

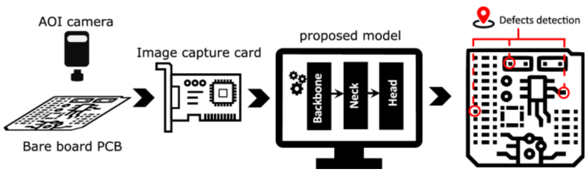

Figure 2. CNN-Based Pipeline for Automated Visual Inspection in PCB Manufacturing

XIV. AUTOMOTIVE AND MECHANICAL COMPONENT INSPECTION

In this industry, CNNs help inspect important engine parts, brake discs and shafts. Because of their typical, irregular surface features such as rust, oil stains or microfractures, these components are often missed by traditional tools for measuring shape. Due to being trained on labelled component photos, CNNs can accurately find and identify defects on a small and large scale (Zheng et al., 2018). Besides, CNNs can handle several imaging modalities—visible, thermal and X-ray—which makes defect detection easier in many different inspection environments. Thanks to this flexibility, the software is better equipped to discover flaws in both easy and hard-to-spot forms under all sorts of lighting conditions.

XV. TEXTILE AND SURFACE PATTERN INSPECTION

Visual inspection of textile products can be tricky because different types of fabric and defects look different each time. Broken weaves, misprints or knots in a carpet usually occur in only one area and can merge with detailed designs. It seems that classical approaches, like Gabor filters and co-occurrence matrices, are not fully effective for representing these different variations (Wang et al., 2018). Standard texture patterns are modeled and anomalies are caught in images with CNN-based unsupervised methods, mainly autoencoders. Li et al. (2018) showed how a convolutional autoencoder can find yarn skips and oil stains in fabric without using any labeled data. Now, these systems are used in monitoring looms to inspect them live and find mechanical faults.

XVI. PHARMACEUTICAL AND PACKAGING QUALITY CONTROL

To guarantee that packages remain intact and follow regulations, inspecting by eye is very important in making drugs. Using CNNs, it is possible to detect missing tablets, misaligned labels and bent blister packs with greater success than is possible with older systems. Because they are good at understanding the relationship between objects, they are used to ensure that expiry dates and tamper proof marks are intact. CNNs have been shown by Nagahara et al., (2017) to discover low-contrast prints on glossy surfaces quicker, reducing the risk of missing anything and enhancing overall flow of items. According to this study, CNNs should be used more widely in areas where finding faults depends on small changes and varies depending on the scene.

XVI. INDUSTRIAL ROBOTICS AND ASSEMBLY LINES

Component positioning, orientation and completeness are verified in industrial robotics using visual inspection systems driven by CNNs. Assemblers use CNN technology to ensure high-precision components have correct screws, adhesive and flawless connections. The quick response time from these systems allows for successful high-speed production. Using transfer learning, CNN models have made it possible for companies to quickly adapt to new products with just a little bit of tagged data. Such systems become very useful in flexible manufacturing, as the variation in products demands instant and flexible inspection (Yamashita et al., 2018).

Industry Sector	Common Defect Types	CNN Methods Used	Benefits
Semiconductor & PCB	Shorts, missing pads, microcracks	ResNet, VGG, CNN classifiers	High-resolution inspection, AOI integration
Automotive Components	Casting cracks, wear, deformation	CNNs with X-ray/thermal imaging	Robust under multimodal conditions
Textile & Fabrics	Weaving faults, print errors, oil stains	Autoencoders, GANs	Label-free learning, texture modeling
Pharmaceuticals	Label misprints, blister deformation	CNN classifiers, OCR CNNs	Regulatory compliance, high sensitivity
Industrial Robotics	Misalignment, missing components	Transfer learning, hybrid CNNs	Fast adaptation, real-time feedback

Table 3. CNN-Based Anomaly Detection Applications across Industrial Sectors

XVII. PERFORMANCE METRICS AND BENCHMARKING

Assessing the ability of CNN-based anomaly detectors in industrial visual inspection is necessary for academic and practical applications. Given that quality control in manufacturing is vital, any assessment of performance must cover how well it finds flaws, how sensitive it is to rare problems, how well it handles actual variability and its speed of computation. As a result, researchers have introduced a variety of evaluation tools and benchmarking data to connect with the distinct needs of diverse applications (Ando et al., 2016; Jabez & Muthukumar, 2015)

XVIII. COMMON EVALUATION METRICS

Even though things such as accuracy, precision, recall and F1-score are used for traditional classification in supervised learning, they prove less valuable when data is highly imbalanced. So, in situations where the ratio of errors is very low, a simple model may do well at predicting the main category but miss all the errors. The Receiver Operating Characteristic Area under the Curve (ROC-AUC) and the Precision-Recall Area under the Curve (PR-AUC) are suggested over other types of metrics for this case. Ge et al. (2017) pointed out that PR-AUC is valuable on unbalanced datasets because it measures the ratio of correct detection of rare anomalies. The IoU measure is vital for defect localization because it counts how much the model's predicted defect regions and the ground truth regions overlap. Also, real-time inspection tools for industrial use depend on data about how much time it takes to process an image and how much memory it uses, as these numbers have direct impact on whether the system can integrate with production machinery.

XIX. BENCHMARK DATASETS FOR INDUSTRIAL INSPECTION

Analyzing different CNN-based anomaly detection models is only possible through the use of benchmark datasets. The DAGM 2007 dataset (Stentoumis et al., 2016) was one of the first and has been most frequently applied in this research area. The syllabus offers grayscale images with synthetically generated problems and clear notes, so it's convenient for early determination of how well a system performs, even though it misses the real-time complexity of industrial scenarios. The Kolektor Surface-Defect Dataset (KolektorSDD) gives a more realistic view of the evaluation process. As explained by Zhao et al. (2017), industrial component images present on this dataset are detailed, allowing models to be assessed by both human experts and by themselves without manual labeling. Li and colleagues put forward the Magnetic Tile Defect Dataset (2018) which contains samples of different defects, for example, cracks, breaks and irregular patterns on magnetic surfaces. Although these datasets have supported the benchmarking process, a major problem is that standardized datasets for various domains and imaging types such as X-ray and thermal, are not widely available. Because of industrial secrecy, important datasets in PCB inspection or automotive parts remain unavailable to people outside the company for public benchmarking.

XX. CROSS-METHOD COMPARISONS AND MODEL VALIDATION

Comparing anomaly detection models using convolutional neural networks gets complicated due to the differences in datasets, ways of preparing data and image augmentation used. In the end, some carefully planned evaluations have shed light on these two systems. Sørensen et al. (2017) explored classical autoencoders, variational autoencoders and one-class SVMs using deep features and found that models using deep learning normally perform better than shallow learning and unsupervised learning on most industrial data sets. A lot of researchers use transfer learning as a basic method to review the performance of algorithms. Networks trained with ImageNet (Achlioptas et al., 2009) as data are adjusted on industrial defect data and often handle unseen defect scenarios well using just a few samples. The success can be explained by the fact that low-level visual features stay important in different domains. Table 4 outlines the main metrics used and the popular datasets involved in testing CNN-based anomaly detection models in industry visual inspection.

Metric	Description	Appropriate Context
Accuracy	Overall correctness of predictions	Balanced datasets (rare in industry)
Precision & Recall	Precision = TP/(TP+FP), Recall = TP/(TP+FN)	Anomaly presence estimation
F1-Score	Harmonic mean of precision and recall	Overall detection quality
ROC-AUC	Trade-off between sensitivity and specificity	Binary classification (threshold-free)
PR-AUC	Precision-recall trade-off, better for imbalanced data	Sparse anomaly settings
IoU (Jaccard Index)	Region overlap for pixel-wise localization	Surface anomaly segmentation
Inference Time	Time per sample inference in milliseconds	Real-time deployment validation
DAGM 2007	Synthetic textured surface images with defects	Texture modeling, autoencoders
KolektorSDD	Real industrial surface defects with annotations	Binary/multi-label classification
MT Defect Dataset	Magnetic tile surface cracks and breaks	Small-scale surface anomaly detection

Table 4. Performance Metrics and Benchmark Datasets in Industrial Visual Inspection

6. Challenges and Limitations

Makhzani & Frey (2017) point out that despite significant progress in CNN-based anomaly detection for checking industry equipment, there are several barriers that hold back their use and success. Identifying and managing these challenges will guide future research and how things are implemented. An important obstacle is that even in industrial sectors, defect data is often limited and unevenly spread. Han et al. point out that, since defective samples make up less than 1% of data, supervised CNNs are not well trained because the network requires many diverse examples to identify useful features. Since defects can differ a great deal in shape, size, appearance and texture, compiling a variety of labeled datasets is both a long and expensive process. Shifting and hard-to-notice challenges linked to industrial defects cause an additional problem. Distortions in texture, changes in color or micro-cracks that appear could be easily confused with harmless light, dirt or material issues (Zhao et al., 2018). This variation in the environment can increase the number of false positives and false negatives in CNN-based detectors, mainly since taking pictures is not standardized in real-world settings. Being able to explain what a model does is still a big problem. Because CNNs are not simple and do not reveal how they make decisions, they are frequently called “black boxes.” (Sung et al., 2018) When results from industrial inspections shape the safety and compliance of goods, explainability is usually necessary rather than just a preference. While CAM and saliency maps have helped provide understanding of how a model functions, they are not detailed enough for strict industries (Kos et al., 2018).

Working with deep CNNs to spot small defects using high-resolution images leads to high computation costs. As a result, the models created are generally bulky and take up a lot of resources, making it hard to use them in live or on-device applications. Though GPUs and FPGAs have improved things, efficient architectures are required more than ever. Pruning, quantization and knowledge distillation appear to be good methods for simplifying models and keeping their performance unchanged (Han et al., 2018). In addition, the discipline does not have enough standard, free datasets that accurately reflect the variety found in real production environments. Several existing datasets contain artificial data and are privately owned, making it tough for research to be repeatable and hard to compare methods consistently (Kim et al., 2018). CNNs usually have challenges applying their knowledge to a variety of products, materials or imaging methods. A model may not work the same in a new context because there are differences in texture, shape or the equipment used to sense the environment. Transfer learning and domain adaptation are effective ways to solve the problem, though they often need a bit of labeled data from the target domain and must be well-tuned to work fully (Jabez & Muthukumar, 2015).

XXI. FUTURE DIRECTIONS AND INNOVATIONS

Exciting outcomes from using CNNs for anomaly detection in industry have motivated researchers to find ways to remove current obstacles and let the technology be used in more areas. The area given the most attention is the development of data-saving learning approaches. With limited availability of examples of defects, approaches using little labeled data such as few-shot learning, semi-supervised learning and self-supervised learning, are becoming more popular (Wang et al., 2018). When using few-shot, CNNs only need a little amount of information to function well and through self-supervision, they can teach themselves to recognize items by studying a large number of normal pictures. With the right plans, it becomes possible to considerably lessen the amount of manual annotation required. An additional interesting line of research is to combine generative models with CNNs. Basically, using techniques such as VAEs and GANs, we are able to compare against the standard patterns, highlighting any dissimilarities as anomalies (Kermany et al., 2018). Without much supervision, these frameworks require less annotation and can process fresh manufacturing issues more easily.

Hybrid architectures create new chances for people who need them. When CNNs are added to graph neural networks or attention mechanisms, the results have shown promise in boosting both reliability and interpretability (Sørensen et al., 2018). In particular, attention-based methods can increase a network's ability to perform well by selecting important parts of the image and soothing the noise. For industrial inspection in real time, CNNs must be optimized for use on the edge and embedded systems. The use of pruning, quantization and distilling knowledge makes it easier to deploy efficient models, even if there are strict rules on latency and computer resources. (Han et al., 2018). The need for explainability remains at the heart of research. Further developments should make CNNs more interpretable, delivering key points to those using them. Including confidence scores and logical reasons in explanations could help raise confidence and support certification in important industries (Srivastava et al., 2017). Furthermore, it is important for industry and academia to collaborate in order to supply huge, realistic datasets for industry use. Anomaly detection models could be developed and tested against new standards, as a result of datasets including different defect types, materials, imaging technologies and operating conditions (Tawara et al., 2018). Improving the ability of models to work in various domains is still a major task. Lately, some methods in domain adaptation and meta-learning have focused on models that need very little retraining after they encounter new products, types of defects or imaging configurations (Zhou & Tuzel, 2018). Because frequent product changes are common in flexible manufacturing, these capabilities are very useful.

CONCLUSION

This article provided a thorough exploration of how CNN-based methods are utilized for anomaly detection in industrial visual inspection, outlining key principles and practical implementations. CNNs' remarkable ability to automatically extract meaningful and diverse features from complex images has significantly expanded the range of defects that can be detected—many of which were previously difficult to identify using traditional or rule-based methods. The versatility of CNN models is evident across various industries, including semiconductor manufacturing, automotive parts inspection, textile quality control, pharmaceutical packaging, and robotic assembly.

The integration of CNNs into automated optical inspection and multimodal imaging systems has led to improved detection accuracy, greater processing capacity, and more streamlined inspection workflows. However, the article also highlights persistent challenges such as the limited availability of defective samples, the complexity and variability of anomalies, interpretability issues, and the high computational demands of deep learning models.

These limitations underscore the need for continued research in areas like data-efficient learning techniques, explainable AI, hardware acceleration, and the development of industrially relevant datasets. Emerging solutions—such as few-shot learning, self-supervised methods, generative modeling, hybrid architectures, and domain adaptation—are promising pathways to expanding the applicability of CNN-based anomaly detection in more dynamic and complex environments. Moreover, collaborative efforts to standardize datasets and benchmarks will enhance reproducibility and facilitate meaningful comparisons across studies.

REFERENCES

[1] Achlioptas, P., Diamanti, O., Mitliagkas, I., & Guibas, L. (2017). Representation Learning and Adversarial Generation of 3D Point Clouds. 35th International Conference on Machine Learning, ICML 2018, 1, 67–85. Retrieved from https://arxiv.org/abs/1707.02392v1

[2] Ando, H., Niitsu, Y., Hirasawa, M., Teduka, H., & Yajima, M. (2016). Improvements of classification accuracy of film defects by using GPU-accelerated image processing and machine learning frameworks. In Proceedings - NICOGRAPH International 2016, NicoInt 2016 (pp. 83–87). Institute of Electrical and

Electronics Engineers Inc. https://doi.org/10.1109/NicoInt.2016.15

[3] Assendorp, J. P., & Deep. (2017). Deep learning for anomaly detection in multivariate time series data. The Ninth International Conferences on Pervasive Patterns and Applications Defect 2017, 13(1), 1–12. Retrieved from https://www.metaljournal.com.ua/assets/Journal/english-edition/MMI_2015_7/045

[4] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

[5] Christiansen, P., Nielsen, L. N., Steen, K. A., Jørgensen, R. N., & Karstoft, H. (2016). DeepAnomaly: Combining background subtraction and deep learning for detecting obstacles and anomalies in an agricultural field. Sensors (Switzerland), 16(11). https://doi.org/10.3390/s16111904

[6] Ge, R., Zhu, Y., Xiao, Y., & Chen, Z. (2017). The subway pantograph detection using modified faster R-CNN. In Communications in Computer and Information Science (Vol. 685, pp. 197–204). Springer Verlag. https://doi.org/10.1007/978-981-10-4211-9_20

[7] Han, Z., Wei, B., Mercado, A., Leung, S., & Li, S. (2018). Spine-GAN: Semantic segmentation of multiple spinal structures. Medical Image Analysis, 50, 23–35. https://doi.org/10.1016/j.media.2018.08.005

[8] Jabez, J., & Muthukumar, B. (2015). Detection Approach. Procedia - Procedia Computer Science, 48(Iccc), 338–346. Retrieved from http://dx.doi.org/10.1016/j.procs.2015.04.191

[9] Kaushik, P., & Jain, M. (2018). Design of low power CMOS low pass filter for biomedical application. International Journal of Electrical Engineering & Technology (IJEET), 9(5).

[10] Kaushik, P., Jain, M., & Jain, A. (2018). A pixel-based digital medical images protection using genetic algorithm. International Journal of Electronics and Communication Engineering, 31-37. http://www.irphouse.com/ijece18/ijecev11n1_05.pdf

[11] Kaushik, P., Jain, M., & Shah, A. (2018). A Low Power Low Voltage CMOS Based Operational Transconductance Amplifier for Biomedical Application. https://ijsetr.com/uploads/136245IJSETR17012-283.pdf

[12] Kaushik, P., & Jain, M. A Low Power SRAM Cell for High Speed Applications Using 90nm Technology. Csjournals. Com, 10. https://www.csjournals.com/IJEE/PDF10-2/66.%20Puneet.pdf

[13] Kaushik, P., & Jain, M. (2018). Design of low power CMOS low pass filter for biomedical application. International Journal of Electrical Engineering & Technology (IJEET), 9(5).

[14] Kaushik, P., Jain, M., & Jain, A. (2018). A pixel-based digital medical images protection using genetic algorithm. International Journal of Electronics and Communication Engineering, 31-37. http://www.irphouse.com/ijece18/ijecev11n1_05.pdf

[15] Kaushik, P., Jain, M., & Shah, A. (2018). A Low Power Low Voltage CMOS Based Operational Transconductance Amplifier for Biomedical Application. https://ijsetr.com/uploads/136245IJSETR17012-283.pdf

[16] Kaushik, P., & Jain, M. A Low Power SRAM Cell for High Speed Applications Using 90nm Technology. Csjournals. Com, 10. https://www.csjournals.com/IJEE/PDF10-2/66.%20Puneet.pdf

[17] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell, 172(5), 1122-

1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

[18] Kim, J. Y., Bu, S. J., & Cho, S. B. (2017). Malware detection using deep transferred generative adversarial networks. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 10634 LNCS, pp. 556–564). Springer Verlag. https://doi.org/10.1007/978-3-319-70087-8_58

[19] Kim, J. Y., Bu, S. J., & Cho, S. B. (2018). Zero-day malware detection using transferred generative adversarial networks based on deep autoencoders. Information Sciences, 460–461, 83–102. https://doi.org/10.1016/j.ins.2018.04.092

[20] Kos, J., Fischer, I., & Song, D. (2018). Adversarial examples for generative models. In Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018 (pp. 36–42). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/SPW.2018.00014

[21] Larsen, A. B. L., Sønderby, S. K., Larochelle, H., & Winther, O. (2016). Autoencoding beyond pixels using a learned similarity metric. In 33rd International Conference on Machine Learning, ICML 2016 (Vol. 4, pp. 2341–2349). International Machine Learning Society (IMLS).

[22] Larsen, A. B. L., Sønderby, S. K., Larochelle, H., & Winther, O. (2016). Vae/Gan. Icml, 4, 2341–2349.

[23] Li, W., Wu, G., & Du, Q. (2017). Transferred Deep Learning for Anomaly Detection in Hyperspectral Imagery. IEEE Geoscience and Remote Sensing Letters, 14(5), 597–601. https://doi.org/10.1109/LGRS.2017.2657818

[24] Li, Y., Min, M. R., Shen, D., Carlson, D., & Carin, L. (2018). Video generation from text. In 32nd AAAI Conference on Artificial Intelligence, AAAI 2018 (pp. 7065–7072). AAAI press. https://doi.org/10.1609/aaai.v32i1.12233

[25] Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2018). Towards deep learning models resistant to adversarial attacks. In 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings. International Conference on Learning Representations, ICLR.

[26] Makhzani, A., & Frey, B. (2017). PixelGAN autoencoders. In Advances in Neural Information Processing Systems (Vol. 2017-December, pp. 1976–1986). Neural information processing systems foundation.

[27] Nagahara, H., Umeda, K., & Yamashita, A. (2017). Thirteenth international conference on quality control by artificial vision 2017. In Proceedings of SPIE - The International Society for Optical Engineering (Vol. 10338). https://doi.org/10.1117/12.2277475

[28] Puneet Kaushik, Mohit Jain , Gayatri Patidar, Paradayil Rhea Eapen, Chandra Prabha Sharma (2018). Smart Floor Cleaning Robot Using Android. International Journal of Electronics Engineering. https://www.csjournals.com/IJEE/PDF10-2/64.%20Puneet.pdf

[29] Puneet Kaushik, Mohit Jain , Gayatri Patidar, Paradayil Rhea Eapen, Chandra Prabha Sharma (2018). Smart Floor Cleaning Robot Using Android. International Journal of Electronics Engineering. https://www.csjournals.com/IJEE/PDF10-2/64.%20Puneet.pdf

[30] Sabokrou, M., Fayyaz, M., Fathy, M., & Klette, R. (2017). Deep-Cascade: Cascading 3D Deep Neural Networks for Fast Anomaly Detection and Localization in Crowded Scenes. IEEE Transactions on Image Processing, 26(4), 1992–2004. https://doi.org/10.1109/TIP.2017.2670780

[31] Sørensen, R. A., Rasmussen, J., Nielsen, J., & Jørgensen, R. N. (2017). Thistle detection using convolutional neural networks: EFITA 2017 Presentation. EFITA WCCA Congress. 2.-6. June 2017, 16(July), 1–15. Retrieved from https://www.mdpi.com/165332

[32] Srivastava, A., Valkov, L., Russell, C., Gutmann, M. U., & Sutton, C. (2017). VEEGAN: Reducing mode collapse in GANs using implicit variational learning. In Advances in Neural Information Processing Systems (Vol. 2017-December, pp. 3309–3319). Neural information processing systems foundation.

[33] Stentoumis, C., Protopapadakis, E., Doulamis, A., & Doulamis, N. (2016). A holistic approach for inspection of civil infrastructures based on computer vision techniques. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives (Vol. 41, pp. 131–138). International Society for Photogrammetry and Remote Sensing. https://doi.org/10.5194/isprsarchives-XLI-B5-131-2016

[34] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H. S., & Hospedales, T. M. (2018). Learning to Compare: Relation Network for Few-Shot Learning. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 1199–1208). IEEE Computer Society. https://doi.org/10.1109/CVPR.2018.00131

[35] T., J. (2018). Comparative Study of GAN and VAE. International Journal of Computer Applications, 182(22), 1–5. https://doi.org/10.5120/ijca2018918039

[36] Tawara, N., Kobayashi, T., Fujieda, M., Katagiri, K., Yazu, T., & Ogawa, T. (2018). Adversarial autoencoder for reducing nonlinear distortion. In 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2018 - Proceedings (pp. 1669–1673). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.23919/APSIPA.2018.8659540

[37] V.N., E., & U.R., I. (2016). Modeling approaches for time series forecasting and anomaly detection. Animal Molecular Breeding, 8–13.

[38] Wang, H., Qin, Z., & Wan, T. (2018). Text generation based on generative adversarial nets with latent variables. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 10938 LNAI, pp. 92–103). Springer Verlag. https://doi.org/10.1007/978-3-319-93037-4_8

[39] Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018, August 1). Convolutional neural networks: an overview and application in radiology. Insights into Imaging. Springer Verlag. https://doi.org/10.1007/s13244-018-0639-9

[40] Zhao, Y., Deng, B., Huang, J., Lu, H., & Hua, X. S. (2017). Stylized adversarial autoencoder for image generation. In MM 2017 - Proceedings of the 2017 ACM Multimedia Conference (pp. 244–251). Association for Computing Machinery, Inc. https://doi.org/10.1145/3123266.3123450

[41] Zhao, Z., Li, B., Dong, R., & Zhao, P. (2018). A surface defect detection method based on positive samples. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 11013 LNAI, pp. 473–481). Springer Verlag. https://doi.org/10.1007/978-3-319-97310-4_54

[42] Zheng, Y. J., Zhou, X. H., Sheng, W. G., Xue, Y., & Chen, S. Y. (2018). Generative adversarial network based telecom fraud detection at the receiving bank. Neural Networks, 102, 78–86. https://doi.org/10.1016/j.neunet.2018.02.015

[43] Zhou, Y., & Tuzel, O. (2018). VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 4490–4499). IEEE Computer Society. https://doi.org/10.1109/CVPR.2018.00472

[44] Puneet Kaushik, Mohit Jain, Study and Analysis of Image Encryption Algorithm Based on Arnold Transformation.

International Journal of Computer Engineering and Technology, 9(5), 2018, pp. 59-63. http://iaeme.com/Home/issue/IJCET?Volume=9&Issue=5