

Deep Learning and Image Processing Algorithms for Detecting and Classifying Surface and Structural Defects in Metal Components: A Systematic Review

Abstract

The function of deep learning and image processing algorithms in identifying and categorizing structural and surface flaws in metal components is examined in this systematic review. The study, which was carried out in accordance with PRISMA principles, finds and examines 51 pertinent publications from the Web of Science and Scopus databases that were published between 2015 and 2024. The review focuses on combining several image processing methods with deep learning models, such as transformers, convolutional neural networks (CNNs), and hybrid architectures. The results are divided into seven major categories, such as sensor fusion techniques, data preparation, transfer learning, real-time monitoring systems, and hybrid deep learning architectures. These topics include innovations including multi-scale processing and the usage of pre-trained networks, which represent the changing environment of defect detection in industrial applications. Significant gains in detection accuracy and flexibility have been demonstrated by hybrid architectures, and transfer learning has helped to solve the problem of data scarcity. Challenges including the computational complexity of hybrid models, the requirement for consistent data preparation, and the integration of multi-modal sensor data are also identified in the study. In order to enhance real-world deployment and overcome existing restrictions, this paper highlights the increasing significance of AI-driven defect detection systems in improving the production efficiency and quality control of metal components. It also identifies topics for future research.

Keywords: Deep learning, Image processing, Metal defect detection, Transfer learning, Real-time monitoring, Sensor fusion

1. Introduction

For industrial systems in industries like aerospace, automotive, and civil infrastructure to be safe, dependable, and long-lasting, surface and structural flaws in metal components must be identified and categorized. If ignored, flaws such pitting, corrosion, inclusions, and fractures can impair mechanical performance and result in disastrous failures (Wang et al., 2020). Conventional non-destructive testing (NDT) procedures, including as eddy current methods, magnetic particle inspection, and ultrasonic testing, have long been accepted as industry norms. However, these methods are time-consuming, sometimes need manual interpretation, and have scalability issues in high-throughput settings (Dwivedi et al., 2018). Defect detection has been transformed by developments in computer vision and machine learning, especially deep learning (DL), which automates feature extraction and permits real-time analysis of intricate visual data (Cha et al., 2017). The present level of deep learning and image processing methods for detecting and classifying flaws in metal components is summarized in this systematic study, which also highlights the ongoing difficulties and transformational potential of these techniques.

Modern engineering systems depend heavily on metal components, whose structural soundness has a direct bearing on operational safety. For example, under stress, corrosion in pipeline welds or microcracks in aviation turbine blades can spread, resulting in failures with serious human and financial repercussions (Golding et al., 2022). Although they work well in controlled environments, conventional NDT techniques are limited in their ability to identify tiny surface irregularities or subsurface flaws without requiring significant human involvement (Cheng et al, 2018). Additionally, because manual inspection is subjective, there is a chance that human inspectors would miss minor flaws because of weariness or cognitive biases (Jahanshahi & Masri, 2012). These difficulties have sparked interest in automated systems that

improve accuracy and repeatability by fusing computer techniques with high-resolution imaging (Yeum et al., 2015).

The shift from manual to automated workflows for defect identification has been greatly aided by image processing methods. To separate flaws from background textures, early methods included morphological procedures, edge detection, and thresholding (Luo et al., 2020). For instance, by differentiating pixel intensity variations between defective and intact regions, Otsu's approach, a global thresholding technique, has been widely employed to segment fractures in steel surfaces (Neogi et al., 2017). However, these techniques frequently fall short in complicated surface geometries or varied illumination circumstances, where fluctuations in texture and noise can imitate fault patterns (Jeon et al., 2015). Texture analysis methods, such as Gabor filters and gray-level co-occurrence matrices (GLCM), were developed to measure surface abnormalities in order to overcome these constraints (Chen et al., 2021). Although these methods increased resilience, their dependence on manually designed elements limited their capacity to adjust to a variety of fault kinds and materials (Liang et al., 2022).

By automating feature learning straight from raw pixel data, deep learning has solved many of the drawbacks of conventional image processing. Because of its hierarchical architecture, which captures spatial hierarchies in pictures, convolutional neural networks (CNNs) in particular have become a key component of defect detection systems (LeCun et al., 2015). For example, Cha et al. (2017) showed that by learning discriminative characteristics across various dimensions, CNNs might perform better than conventional techniques in spotting concrete fractures. For real-time defect localization on metal surfaces, region-based CNNs (R-CNNs) and You Only Look Once (YOLO) architectures have also been modified, allowing for accurate bounding box predictions around anomalies (Khan et al., 2021).

The incorporation of generative adversarial networks (GANs) to enhance training datasets is a noteworthy advancement in DL-based defect identification. In industrial applications with few labeled defect samples, GANs mitigate data scarcity by synthesizing realistic defect pictures (Tseng et al., 2023). For instance, Kim et al. (2020) improved the generalization of classification models by creating artificial corrosion patterns on aluminum alloys using CycleGAN. With pre-trained models like ResNet and VGG16 optimized for defect-specific tasks, transfer learning has also gained popularity and reduced training time and computational expenses (Miorelli et al., 2022).

Notwithstanding these developments, a number of obstacles prevent DL-based systems from being widely used. First, it might be challenging to achieve consistent detection accuracy due to the variety in defect shape, such as uneven fracture geometries or diffuse corrosion patterns (Yu & Zhang, 2023). According to Wang et al. (2023), DL models frequently need for sizable annotated datasets, which can be difficult to curate, particularly for uncommon or novel fault types. Third, the implementation of state-of-the-art designs, including Vision Transformers (ViTs), in situations with limited resources is limited by their computational complexity (Dosovitskiy et al., 2021). Furthermore, interpretability is an issue due to the "black-box" nature of deep learning models, especially in safety-critical industries where decision transparency is crucial (Samek et al., 2021).

Hybrid approaches that combine DL and conventional image processing to capitalize on their complementing advantages have been studied recently. To improve feature extraction in multi-scale defect analysis, for example, CNNs have been used with wavelet transformations (Sun et al., 2025). Similarly, to increase the accuracy of defect identification in load-bearing components, physics-informed neural networks (PINNs) integrate material science concepts, such as stress distribution models (Raissi et al., 2019). By filling in the gaps in solely data-driven methods, these hybrid frameworks seek to strike a compromise between computational efficiency and domain-specific expertise.

This systematic review evaluates peer-reviewed studies on deep learning and image processing algorithms for metal defect detection, focusing on their technical innovations, limitations, and industrial applicability. By synthesizing findings from diverse applications—such as weld inspection, additive manufacturing, and fatigue crack monitoring—this work identifies trends, benchmarks performance metrics, and highlights unresolved challenges. Furthermore, it proposes directions for future research, emphasizing the need for lightweight architectures, explainable AI, and cross-domain adaptability.

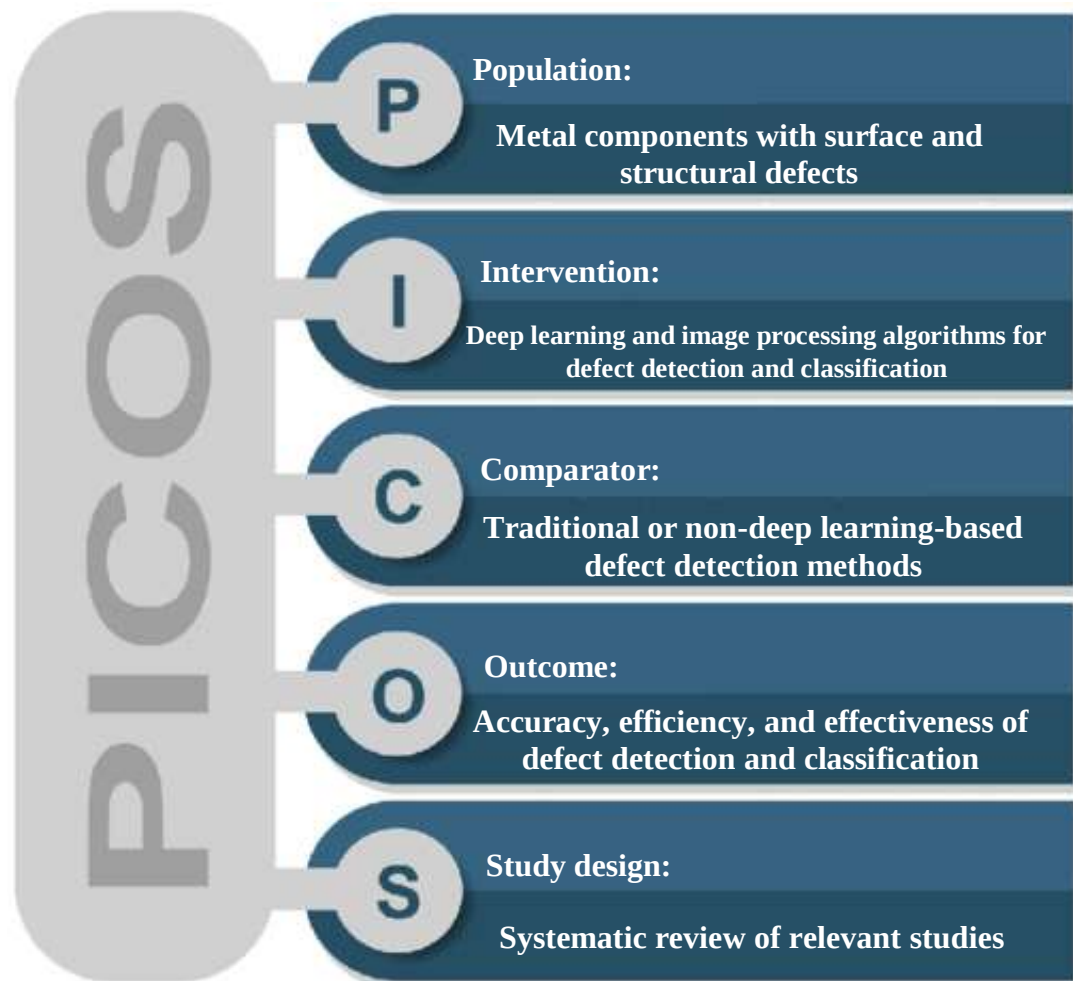


Figure 1. PICOS search terms

To ensure a comprehensive review of the literature, the PICOS framework was used in this investigation. The scope and methods of this systematic study of deep learning and image processing algorithms for identifying and categorizing surface and structural flaws in metal components may be defined in an organized manner using the PICOS framework. A clear focus on industrial applications is ensured by the study's identification of Population (P) as metal components. The Intervention (I) evaluates the function of many deep learning and image processing methods in defect identification. The improvements and possible advantages of AI-driven solutions are demonstrated by the Comparison (C) with conventional techniques. In order to ensure a critical assessment of the evaluated research, the Outcome (O) focuses on important performance criteria including accuracy, efficiency, and dependability. Finally, by methodically examining and integrating the body of existing literature, the Study Design (S) establishes the methodological rigor. Defect identification in industrial metal components

benefits greatly from this systematic method, which improves transparency, repeatability, and the validity of results.

Research Questions

RQ1: What are the most effective deep learning and image processing algorithms used for detecting and classifying surface and structural defects in metal components?

RQ2: How do deep learning-based defect detection methods compare to traditional techniques in terms of accuracy, efficiency, and reliability?

RQ3: What are the key challenges and future research directions in the application of deep learning and image processing for defect detection in metal components?

2. Methods

To guarantee a thorough and open selection procedure, this systematic review was carried out in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standards. The goal of the study was to find and evaluate pertinent studies on deep learning and image processing techniques for identifying and categorizing structural and surface flaws in metal parts.

Database Selection and Search Strategy

We performed a thorough search in the Web of Science (WOS) and Scopus databases to gather pertinent material since these platforms offer a large selection of scholarly publications with a high impact. The following keywords were used in the search: "metal defect" AND "Deep Learning" or "Image Processing" to find publications published between 2015 and 2024.

Only 16 of the 31 documents that came up in the Scopus search were research articles; the other documents were conference papers, reviews, and proceedings. We found 115 papers in Web of Science (WOS), 94 of which were classified as articles. We moved on to the screening and selection procedure after eliminating publications that weren't articles.

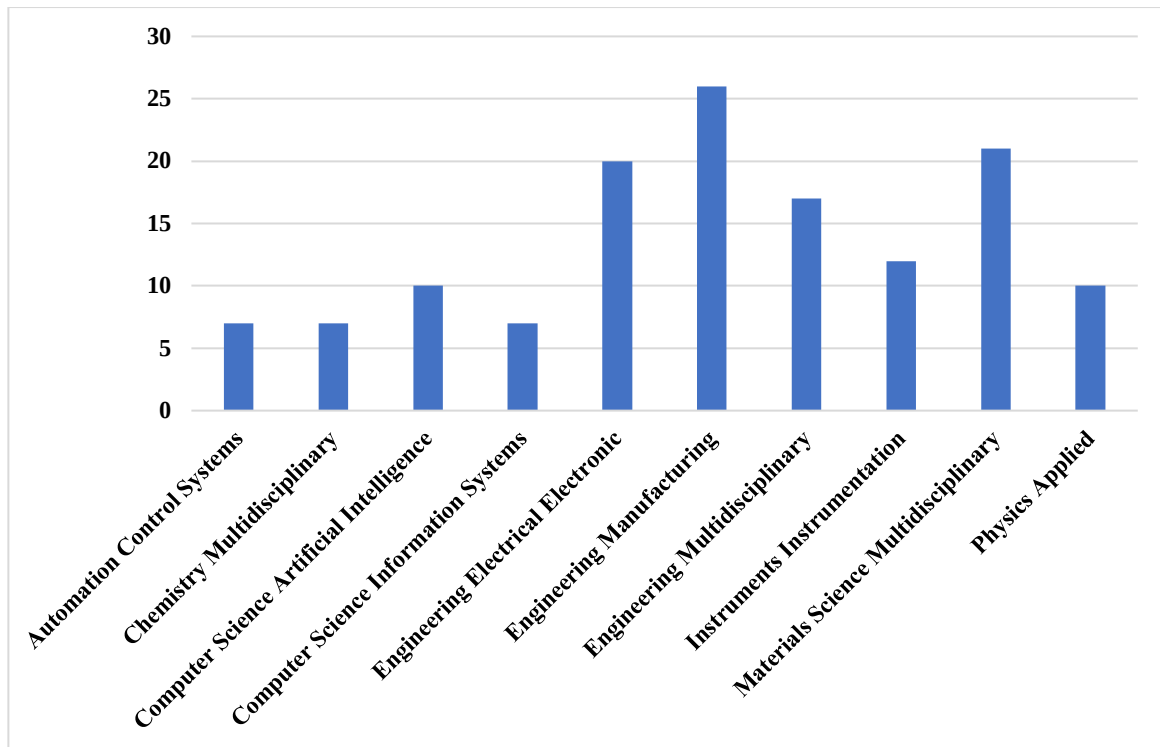


Figure 2: Number of Documents in WOS based on Category

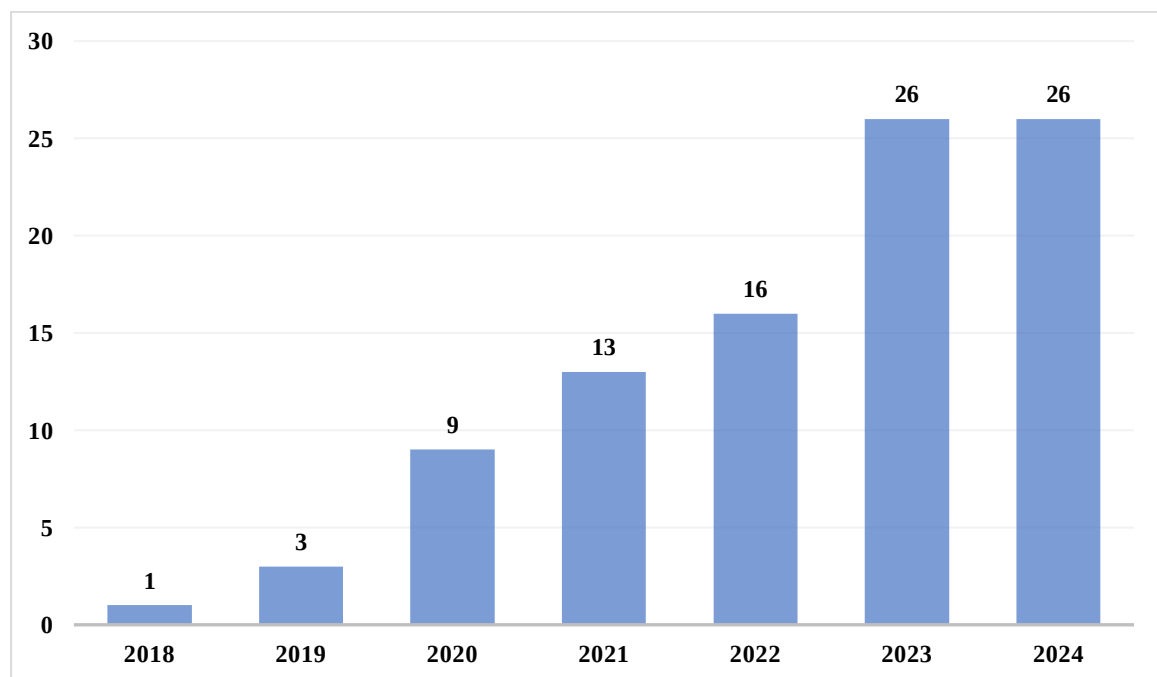


Figure 3: Number of Documents in WOS based on Publication Year

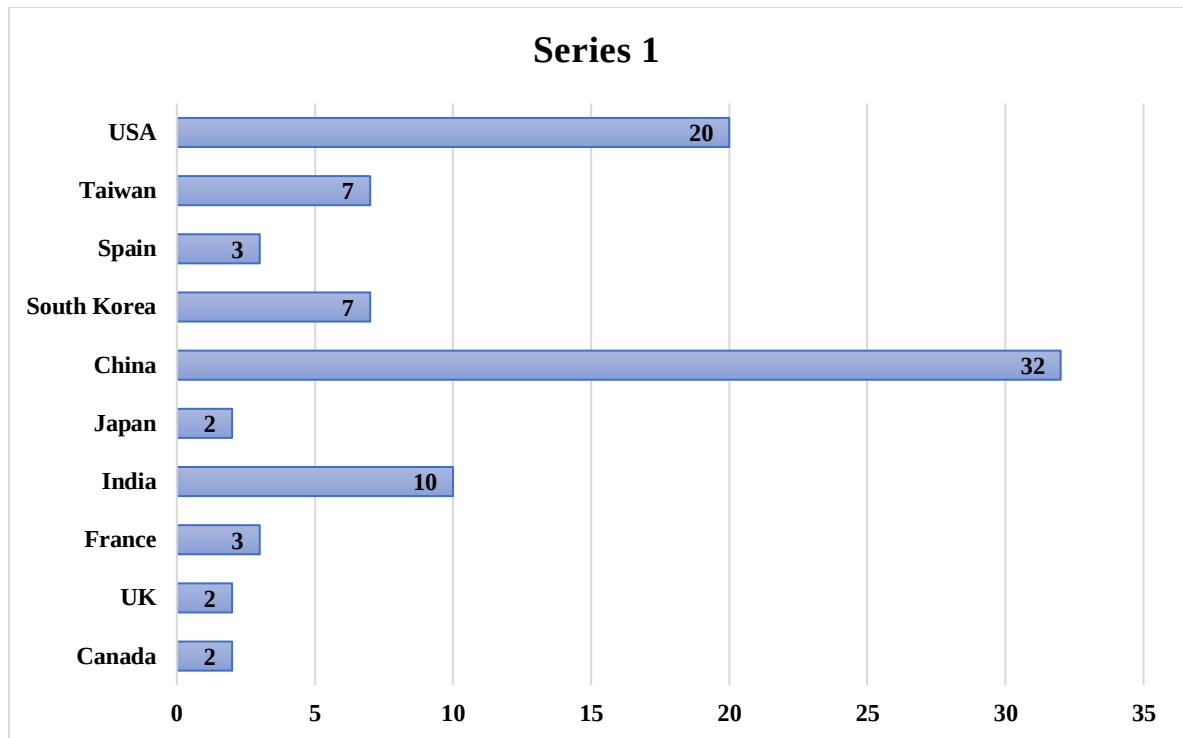


Figure 4: Number of Documents in WOS based on Country

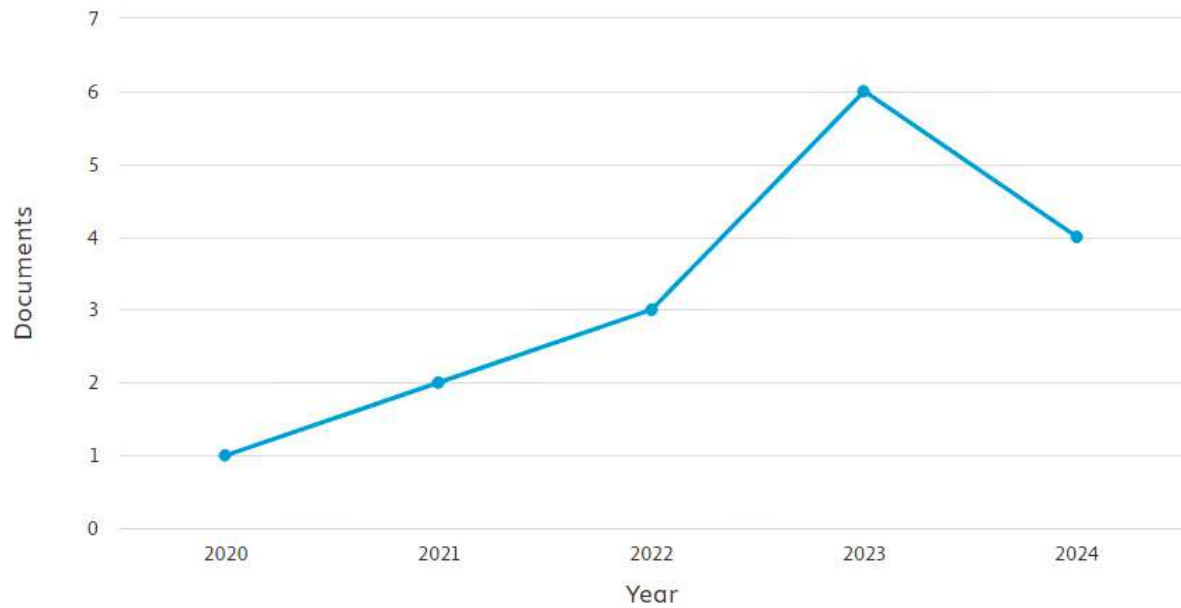


Figure 5: Number of Documents in Scopus based on Publication Year

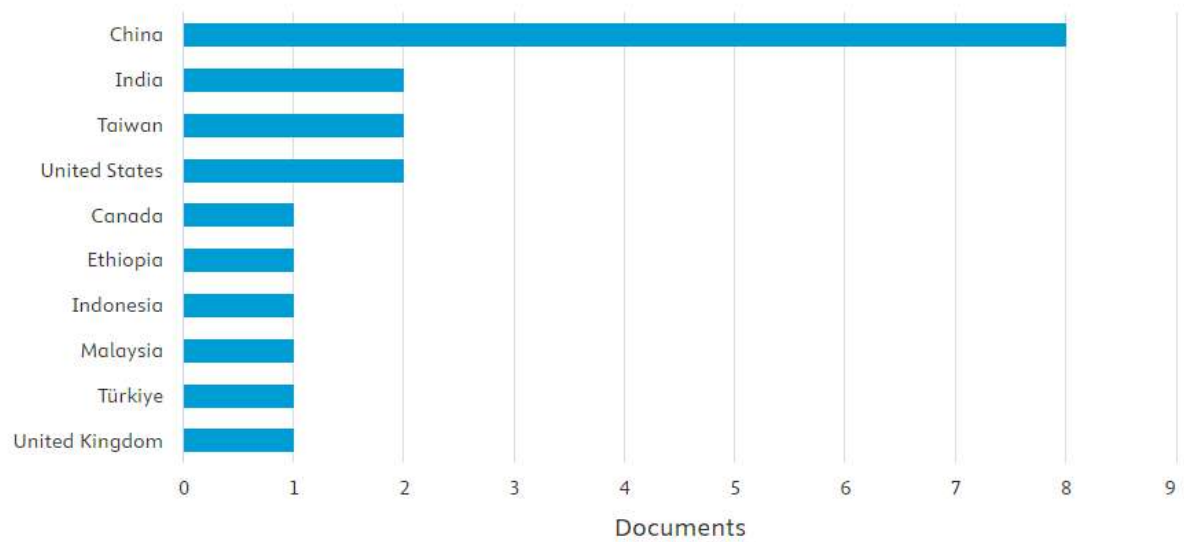


Figure 6: Number of Documents in Scopus based on Country

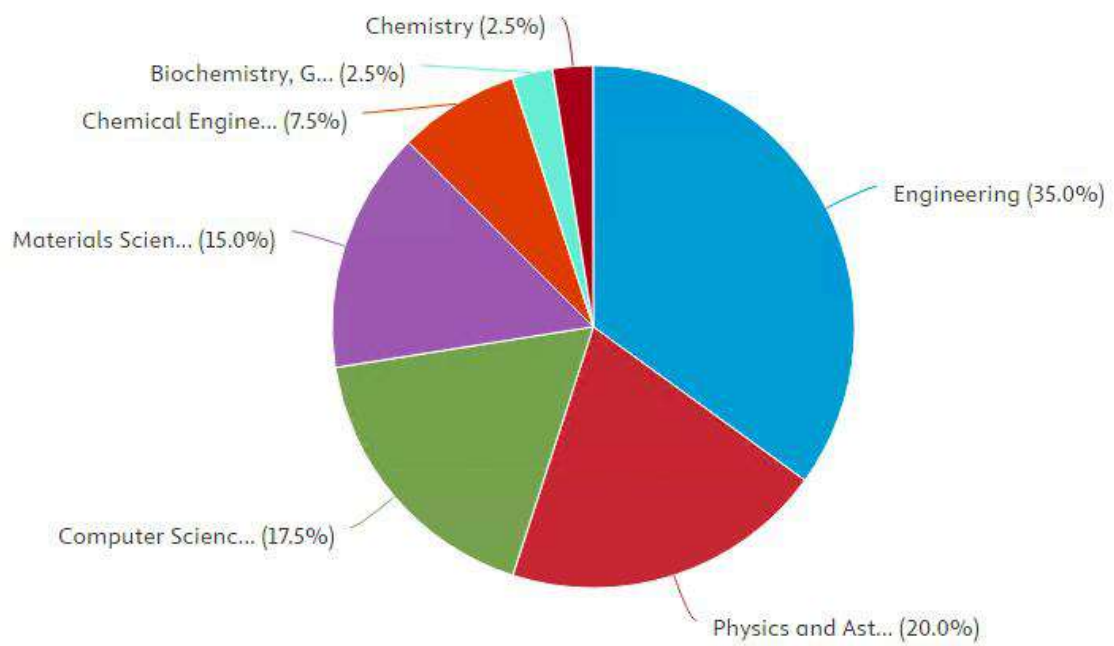


Figure 7: Number of Documents in Scopus based on Category

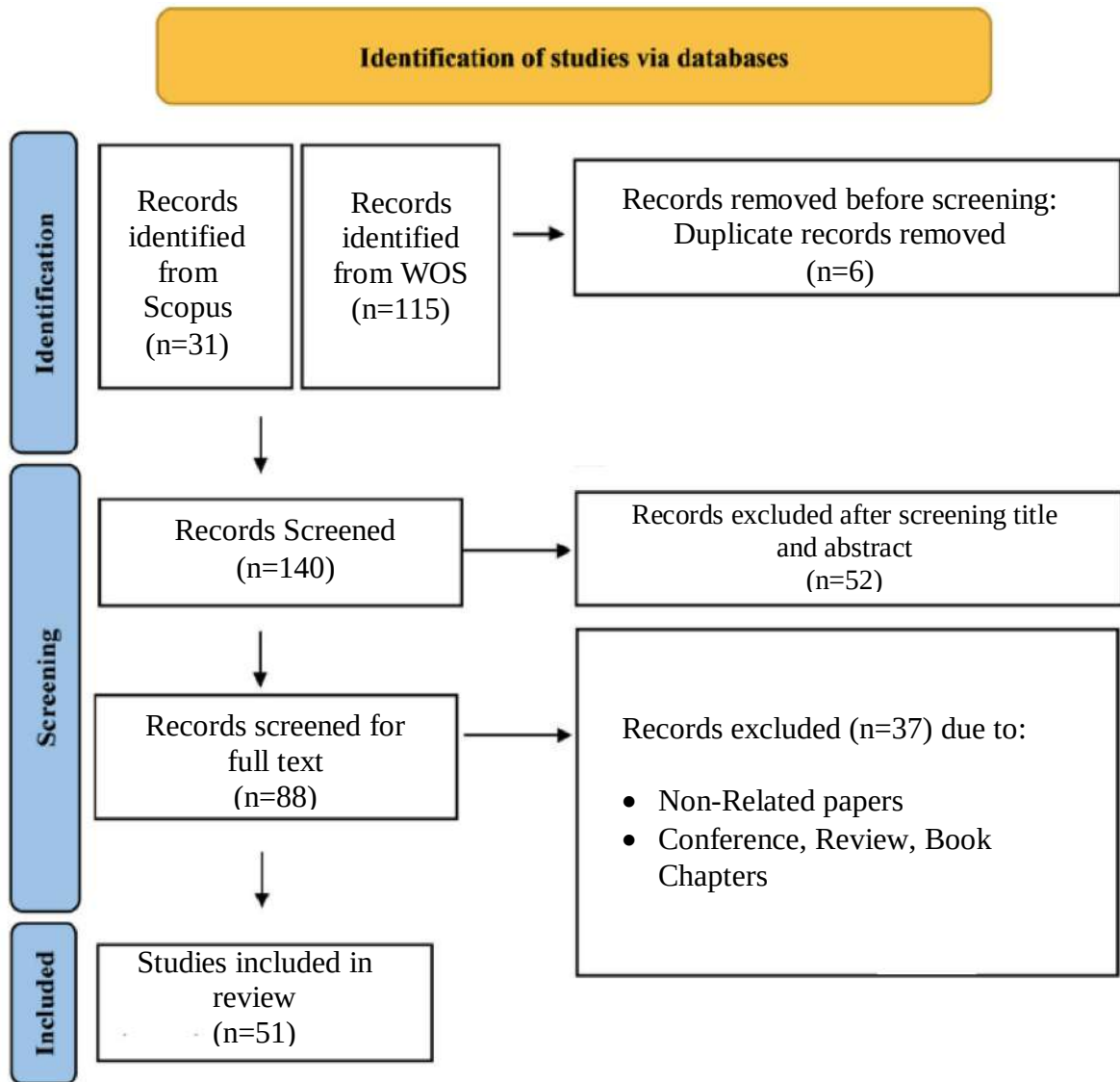


Figure 8. PRISMA Flow Diagram

Inclusion and Exclusion Criteria

The selection process involved a two-step screening:

Title and abstract screening to remove obviously irrelevant studies.

Full-text assessment to ensure the relevance of the research to our objectives.

The following exclusion criteria were applied:

Papers unrelated to our research questions.

Conference papers, review papers, and proceedings.

Studies that did not focus on deep learning or image processing for metal defect detection.

Following this process, 51 final articles were selected for in-depth analysis. Additionally, six papers were identified as duplicates across both databases and removed.

Bibliometric Insights

The most active years for research in this area were 2023 and 2024, according to a review of publication patterns, suggesting a recent uptick in interest and developments in deep learning-based defect identification. As a result of their leadership in industrial AI research, China and the USA contributed the most papers in both databases in terms of geographic distribution.

The fact that Engineering Manufacturing was the most popular topic area in WOS demonstrated how widely these techniques are used in industry.

In Scopus, the Engineering category also had the most publications, further confirming the field's technical focus.

These findings illustrate that deep learning applications in metal defect detection have been a rapidly evolving research area, with significant contributions from engineering and manufacturing domains.

3. Results

In order to identify and categorize surface and structural flaws in metal components, this systematic review examined 51 carefully chosen academic publications on deep learning and image processing techniques. These works were found following a thorough screening procedure based on methodology, relevance, and contribution to the field. They were published between 2015 and 2024. The chosen articles offer a thorough summary of current developments and new trends in a variety of deep learning architectures, image processing methods, and industrial applications.

The main conclusions of these investigations are compiled in Table 1, which also lists the deep learning models that were utilized, the kinds of flaws that were found, the datasets that were used, and the assessment metrics that were provided. Additionally, the table groups the examined articles according to their main areas of interest, including hybrid deep learning models, CNN-based fault classification, comparisons with conventional techniques, and real-time industrial applications. This methodical examination identifies possible topics for further study while highlighting the advantages and disadvantages of the existing strategies.

Table 1: Summary of 51 papers

No	Reference	Aim	Methods	Findings	Contributions
1	Chen et al, 2024	To develop a hybrid detection method that accurately detects and classifies defects in metal pipeline components as a proxy for metal defect inspection.	They combined YOLOv5 with Vision Transformer (ViT) using a cascaded approach and an in-house laboratory dataset generated by Pull Through Testing.	The cascaded approach achieved higher classification accuracy than using YOLOv5 alone while preserving high precision in defect detection.	Introduces an advanced hybrid deep learning model that improves defect detection and classification reliability for metal components, serving as a benchmark for future studies in this area.
2	Guo et al, 2022	To predict transient thermal signatures in laser metal deposition (LMD) processes for early defect detection in metal additive manufacturing.	They developed a deep learning surrogate model (LMD-cGAN) integrated with a physics-guided image selection (PGIS) mechanism to generate thermal signature images conditionally based on deposition layers.	The model effectively predicted melt pool thermal signatures in real time and captured the inherent process uncertainties, proving its robustness.	Provides an efficient, real-time deep learning framework for defect prevention in metal components produced by additive manufacturing, overcoming limitations of traditional physics-based models.
3	Satsangee et al, 2024	To evaluate the performance of the YOLOv7 model for detecting surface defects in metal components, particularly in a remanufacturing context.	They utilized transfer learning on several public metal defect datasets (NEU-DET, NRSD, KolektorSDD2) and conducted an industrial case study to validate model performance.	The tiny variant of YOLOv7 achieved a 73.9% mAP with high FPS, while other variants performed strongly on different datasets, indicating robust real-time detection capabilities.	Validates the application of advanced YOLOv7 deep learning algorithms for efficient, real-time defect inspection in metal surfaces, contributing valuable insights to automated quality control in remanufacturing.
4	Yazdi et al, 2020	To monitor laser powder bed fusion (LPBF) processes for identifying porosity defects in metal components during additive manufacturing.	A hybrid deep neural network was developed using statistical features from wavelet transforms and texture analysis combined with	The proposed framework achieved an F-score of 97.14% in detecting porosity defects, demonstrating robust performance	Provides a real-time inspection framework that enhances the quality assurance process in metal additive manufacturing

			original powder bed images to assess process conditions.	under real-time conditions.	by accurately identifying defects before the final part is completed.
5	Zhang et al, 2024	To enhance laser ultrasonic detection of metal surface defects by integrating digital twin data with deep learning for improved defect recognition.	They used finite element methods to generate twin data, minimized distribution differences with GANs, and applied a residual network with a global attention mechanism to synchronize digital twin models with real equipment.	The proposed method achieved high detection accuracy and demonstrated strong generalization ability, effectively addressing issues related to small sample sizes in defect detection.	Advances digital twin-based inspection methods by combining them with deep learning, providing a scalable solution for accurate defect detection in metal components under industrial conditions.
6	Liang et al, 2024	To improve defect detection in metal stamping processes by refining deep learning algorithms tailored to handle diverse product dimensions and image quality.	An improved YOLOv5 model was employed, enhanced with a BotNet structure and multiheaded self-attention mechanisms, along with image processing modules to reduce reflections and augment data using stable diffusion.	The approach achieved an mAP of 98.2% with reduced model parameters, demonstrating both high accuracy and fast detection speed in challenging industrial conditions.	Offers a state-of-the-art deep learning solution for metal stamping defect detection that addresses variable product characteristics and enhances image quality, thus advancing industrial quality control methods.
7	Wu et al, 2020	To develop an automated defect detection system for metallographic analysis in metal components using advanced convolutional neural network architectures.	A novel CNN architecture called Multi-scale ResNet (M-ResNet) was proposed, incorporating multi-scale operations to improve the detection of small and varying-sized defects in metallographic images.	The modified ResNet achieved an 85.7% mAP in defect recognition, outperforming several existing methods and demonstrating reliable detection capabilities.	Introduces a new deep learning approach for metallographic inspection that improves the efficiency and accuracy of detecting micro-defects, thereby supporting enhanced quality control in metal manufacturing.
8	Sheu et al, 2021	To automate the detection of various defects in sheet metal-based products by integrating multiple imaging angles with deep learning.	A multi-camera system was implemented, using a combination of LeNet and Global Average Pooling (GAP) CNN along with SHAP values for preprocessing to extract and classify image features from different angles.	The method achieved exceptional performance metrics with precision, recall, and AUC nearing 1.00, indicating highly effective defect detection across multiple defect types.	Provides an integrated AI-based inspection solution that significantly improves defect detection in sheet metal components, enhancing quality control through multi-angle image analysis and deep learning techniques.
9	Yan et al, 2020	To create a vision-based inspection system for real-time detection and classification of metal surface defects under imbalanced data conditions.	The study utilized a convolutional variational autoencoder (CVAE) combined with a deep CNN classifier, generating sufficient defect data and overcoming data imbalance issues through conditional data generation.	The proposed method demonstrated excellent performance in classifying defects and effectively handled the imbalanced data problem in a real manufacturing setting.	Introduces a novel deep learning framework that leverages data generation to enhance defect detection in metal surfaces, contributing to more reliable real-time quality control solutions in industrial applications.
10	Yedukondalu et al, 2024	To automatically detect various surface defects in metal components using deep learning models optimized for industrial quality control.	Six different CNN architectures, including ResNet18, were tested with comprehensive image preprocessing, augmentation, and optimization using the Adam optimizer to identify defect types accurately.	ResNet18 emerged as the top performer, achieving an accuracy of 99.77% in defect detection, proving the robustness of the model under varied industrial conditions.	Provides a highly effective deep learning-based solution for automatic defect detection in metal surfaces, demonstrating significant improvements in inspection accuracy and reliability for quality control.
11	Gong et al, 2020	To automate defect detection in small metal cylindrical shells (MCS) used in critical industrial applications through	Transfer learning was employed using Inception-v3 for feature extraction, followed by logistic regression to	The system achieved a 97% accuracy rate, confirming its ability to detect defects with	Offers a practical and efficient deep learning-based detection system for high-precision defect inspection in

		image processing techniques.	classify defects, with five images per MCS analyzed to yield a composite detection outcome.	high precision in metal cylindrical shells.	metal cylindrical components, supporting improved quality assurance in industrial production.
12	Zhang et al, 2023	To develop an enhanced method for detecting scratches on metal flat surfaces by addressing challenges in edge clarity and noise reduction.	An improved Gauss-Laplace (LoG) operator was combined with deep learning-based local texture screening to differentiate scratch edges from interference edges in images of metal surfaces.	The proposed method achieved near-perfect detection rates across multiple scratch categories, effectively distinguishing between true scratches and noise.	Advances the field of edge detection for metal surface inspection by providing a robust method for accurately identifying scratch defects, thus enhancing the overall quality control process in metal components.
13	Vengaloor et al, 2022	To design an automatic quality control system for metal surfaces by accurately segmenting and detecting defect regions through deep learning.	A high-caliber U-shaped deep learning network with up-sampling techniques was implemented for semantic segmentation of defect regions, validated on benchmark datasets such as NEU and Kolektor.	The system demonstrated high accuracy, precision, and Intersection over Union (IoU) metrics, effectively locating and identifying defect regions in metal surfaces.	Provides a robust segmentation-based approach for defect detection that improves automated quality control in metal manufacturing by accurately identifying and localizing surface anomalies.
14	Chandra et al, 2023	To classify regular and irregular bead formations in wire-arc additive manufacturing (WAAM) to detect deposition anomalies in metal components.	A convolutional neural network (CNN) was applied to images captured by a digital camera during the WAAM process, with experiments conducted using different training and testing ratios to optimize performance.	The CNN model achieved an accuracy between 86.53% and 88.08% depending on the number of training epochs, effectively distinguishing between regular and irregular beads.	Contributes a novel deep learning-based classification method for anomaly detection in additive manufacturing processes, thereby improving the quality of deposited layers and mechanical properties of metal parts.
15	Zubayer et al, 2023	To inspect metallographic images for microstructural defects such as cracks and porosity in metal additive manufacturing using advanced deep learning.	The study utilized an enhanced YOLOv8 algorithm with dilated convolution mechanisms for image segmentation and defect detection on a metallography dataset from additive manufacturing of steels.	The proposed approach achieved a defect detection accuracy of up to 96% within a short processing time, demonstrating high efficiency in identifying microstructural defects.	Introduces a novel deep learning framework that significantly improves the detection of microstructural defects in metal components, paving the way for more accurate quality inspections in additive manufacturing.
16	Dubey et al, 2024	To assess how standardized labeling impacts deep learning-based defect detection in metal components.	They evaluated YOLOv5s using confidence intervals on the GC10 DET benchmark and compared precise vs. inconsistent labeling.	Precise labeling with accurate bounding boxes increased average precision by 12.26–25.78%.	Emphasizes the importance of data validation and standardization in training deep learning models for metal defect detection.
17	Xue & Wu, 2022	To develop a deep learning algorithm for detecting and classifying defects in metal nanomaterials.	They implemented a multi-mode detection and classification method using a deep learning network combined with big data analytics.	The study identified five key parameters and achieved quantitative detection of metal nanomaterial defects.	Advances defect detection in metal nanomaterials by integrating deep learning with predictive big data analysis for future trend estimation.
18	Williams et al, 2018	To detect surface and subsurface defects in additively manufactured metal parts during the LPBF process.	They combined Spatially Resolved Acoustic Spectroscopy (SRAS) with a deep learning CNN (DCB-MIR) to convert acoustic maps into optical-like images.	The DCB-MIR model significantly improved the cosine similarity between predicted and actual optical images, enabling better defect identification.	Introduces an innovative approach that integrates acoustic sensing and deep learning for real-time defect detection in metal components produced by LPBF.
19	Zou et al, 2024	To automatically detect and classify defects in metal-ceramic substrates in industrial power device manufacturing.	They utilized a ResNet-50-based CNN combined with multidirectional filter processing and optimized illumination techniques on a	The model achieved a defect discrimination accuracy of 99.8% across various metrics.	Provides a practical deep learning framework for accurate and efficient defect detection in metal-ceramic substrates,

			specialized defect dataset.		supporting industrial quality control.
20	Meng et al, 2021	To evaluate the depth of metal surface defects using an automatic deep learning-based approach.	They developed a portable ECT device and constructed the MDDECT dataset, then trained a 1-D residual CNN (ResNeXt) to classify defect depths as a time series problem.	The 38-layer ResNeXt model achieved a 93.58% accuracy in classifying defect depths from 0.3 to 2.0 mm.	Offers a novel, portable solution for the automatic evaluation of defect depth in metal components, enhancing inspection accuracy and consistency.
21	Kim et al, 2024a	To design a high-speed defect detection system for printed circuit boards (PCBs) in smart manufacturing environments.	They used a line-scan camera for side-view imaging and trained a deep learning model on high dynamic range images of both flawless and defective PCBs.	The system achieved a defect detection accuracy of 99.5% in experimental evaluations.	Provides an efficient high-speed deep learning system for PCB defect inspection, improving quality control in small- and medium-sized enterprises.
22	Seker et al, 2023	To detect welding faults in resistance spot welding processes using deep learning for improved quality in metal component fabrication.	They applied transfer learning networks on a dataset of welding images collected under real production conditions, focusing on the copper-to-copper welding process.	The method achieved classification accuracies ranging from 89% to 100% in detecting welding faults.	Offers an effective deep learning-based solution for weld fault detection, addressing an underexplored area in the welding process and enhancing industrial production quality.
23	Madhavan et al, 2024	To diagnose faults in sheet metal manufacturing by integrating deep learning for feature extraction with machine learning for classification.	They extracted features using six pre-trained CNN models and applied the J48 decision tree for feature selection, followed by various classifiers to identify defects.	The combination of DenseNet201 features with a functional tree classifier achieved an overall accuracy of 99.72%.	Sets a new benchmark in sheet metal fault diagnosis by integrating deep and machine learning techniques, significantly enhancing detection accuracy compared to traditional methods.
24	Yu & Guo, 2024	To segment metal surface defects using a weakly supervised deep learning network with minimal annotation requirements.	They employed a triplet-image input with a weight-shared backbone, an attention activation aggregation module, and a dynamic reasoning module combined with triplet loss.	The proposed method achieved state-of-the-art segmentation performance on six metal surface defect datasets.	Reduces the annotation burden by using image-level labels and weak supervision, thus advancing efficient defect segmentation in metal components.
25	Guo et al, 2024	To segment and detect metal corrosion areas by integrating global and local feature encoding in a deep learning framework.	They proposed MCD-Net, a network that combines a visual Transformer sequence encoder with convolutional encoder-decoder architecture and an attention-based feature fusion module.	The method achieved an F1 score of 84.53% on a public corrosion dataset, demonstrating robust boundary detection under challenging conditions.	Enhances corrosion segmentation accuracy and robustness, providing a valuable tool for detecting corrosion in metal components with complex surface variations.
26	Mustafaev et al, 2023	To recognize surface defects in servo motor housings using a deep learning-based inspection system tailored for small defect detection.	They developed an Inception-CNN model with an integrated inception module and image patching technique, complemented by a GAN (SDG-GAN) for synthetic data generation to balance the dataset.	The model achieved over 99% accuracy and processed 12 images per second in real-time defect inspection.	Provides a real-time, highly accurate deep learning system for defect recognition in servo motor housings, addressing data imbalance with GAN-generated images.
27	Yang et al, 2024	To identify point defects in 2D materials using deep learning-based segmentation and classification techniques.	They implemented a platform that segments unit cells and further divides them into sub-regions for enhanced defect classification, using deep learning for precise defect detection.	The approach quantified point defects accurately in 2H-MoTe ₂ , including Te monovacancies, divacancies, and Mo monovacancies.	Offers an atomic-scale defect analysis method that links defect characteristics with material properties in 2D materials, advancing defect-property correlation research.
28	Wu et al, 2024	To inspect surface defects in IC metal packages	They developed a framework that uses a	The system achieved an error rate of 3.16%	Introduces a novel GAN-based inspection

		using a GAN-based machine vision framework for robust defect evaluation.	GAN template generator, weight mask, adaptive thresholding, and patch-based evaluation to analyze AOI-acquired images.	and a mission rate of 0.89% with an average inspection time of 119.86 ms per sample.	framework that enhances the accuracy and speed of surface defect detection in IC metal packages, improving overall quality control.
29	Singh et al, 2023	To automatically detect and classify defects in sheet metal stamping using advanced image processing and deep learning techniques.	They utilized high dynamic range imaging combined with CNN-based detectors trained on optimally exposed, tone-mapped images of stamped sheet metal parts.	The deep learning model achieved substantially higher accuracy with minimal false positives under challenging industrial lighting conditions.	Provides an innovative framework that leverages high dynamic range imaging to enhance defect detection in sheet metal stamping, offering improved reliability in harsh manufacturing environments.
30	Snyers et al, 2024	To classify process anomalies in Directed Energy Deposition (DED) for metal additive manufacturing using hyperspectral imaging and deep learning.	They combined hyperspectral imaging with a CNN enhanced by Guided GradCAM for visualization, correlating thermal state information with defect classification in real time.	The classification model achieved over 94% accuracy and demonstrated real-time inference capability, linking melt pool temperature distribution with defects.	Paves the way for smart DED process control by integrating hyperspectral imaging with deep learning, thereby improving defect detection and enhancing part quality in metal additive manufacturing.
31	Deng et al, 2021	To improve weld defect detection robustness in industrial images.	Applied median filtering, image enhancement, and an improved deep CNN with transfer learning on weld images.	Achieved a recognition accuracy of 98.75% for weld defect detection.	Enhances reliability and automation for weld defect detection in sheet metal laser welding.
32	Kim et al, 2024b	To inspect internal defects in ceramic-polymer-metal composites.	Employed THz-TDS imaging with an AnoGAN algorithm and data augmentation techniques.	Detected defect regions with an accuracy exceeding 98%.	Enables non-contact, non-destructive inspection of internal defects in composite materials.
33	Block et al, 2024	To automate detection and classification of weld bead defects.	Developed the LoHi-WELD dataset and applied YOLOv7 variants on low and high-resolution weld images.	Lightweight architecture reached up to 0.69 mAP (fine) and 0.77 mAP (coarse).	Provides a public dataset and model benchmarks for robust weld defect classification in industrial settings.
34	TANG et al, 2022	To recognize metal defects with fewer parameters and calculations.	Combined a shallow CNN with a Transformer and SE channel attention module on metal defect images.	Effectively identified defects on small and medium-sized datasets.	Offers a universal method with reduced computational cost for metal defect recognition.
35	Zhao et al, 2023	To detect metal surface defects in a semi-supervised framework.	Utilized a Transformer with pruned multi-scale masked feature fusion and incorporated ViT in a GAN framework.	Demonstrated superior detection of rust, scratches, and other defects on multiple metal datasets.	Advances semi-supervised defect detection with improved training speed and robust feature learning.
36	Wang et al, 2024	To enhance defect detection on metal surfaces using an improved cascade R-CNN.	Improved the backbone network with multiscale feature maps and applied contrast histogram equalization on the NEU-DET dataset.	Achieved an mAP of 0.754, outperforming the standard cascade R-CNN by 9.2%.	Advances metal defect localization and detection with enhanced deep learning techniques.
37	Sudharsan et al, 2024	To detect and size metal defects by fusing ultrasonic and thermographic data.	Combined PAUT and pulsed thermography with a DNN and Bi-MAT algorithm to process fused volumetric data.	Achieved a flaw detection accuracy of 91.46% with precise defect sizing.	Integrates multi-modal inspection for improved defect geometry extraction in metal components.
38	Chen et al, 2023	To inspect surface defects in IC metal packages via a multi-scale framework.	Designed an entirely multi-scale GAN with a transformer decoder, weight masks, adaptive thresholding, and patch-based evaluation.	Achieved 99.3% accuracy with high precision and a speed of 70.9 FPS.	Provides a state-of-the-art semi-supervised inspection framework for IC metal package defect detection.

39	Elleuch et al, 2024	To improve the classification of steel surface defects under ambiguous conditions.	Applied possibility transform based on probability intervals with deep learning methods on NEU-DET and GC10-DET databases.	Achieved classification accuracies close to 100% for most defect types.	Introduces a novel formalism to handle ambiguity and uncertainty in surface defect classification.
40	Jacintha et al, 2023	To detect and identify flaws in plug valves using thermal imaging.	Combined Dye-Penetrating Test with infrared thermography and a fuzzy deep learning algorithm for image processing.	Achieved a detection accuracy of 80.67% for surface flaw identification.	Enhances thermal image-based defect detection in plug valves by integrating fuzzy algorithms with deep learning.
41	Ma et al, 2022	To quantitatively detect defects on complex curved metal surfaces.	Used a flexible eddy current sensor on a four-axis platform and applied maximum interclass variance for 3D defect reconstruction.	Reconstructed 3D defect images with quantitative errors below 6%.	Provides a quantitative method for defect detection on curved metal surfaces with high precision.
42	Yu et al, 2022	To segment metal surface defects with few-shot learning approaches.	Proposed SPNet with a matrix decomposition attention mechanism and conditional boundary refinement.	Achieved state-of-the-art segmentation performance on the Defects dataset.	Improves defect segmentation with limited annotated data using few-shot learning and advanced attention modules.
43	Katsamenis et al, 2022	To simultaneously localize and classify corrosion and rust grades on metal structures.	Employed SPLAC U-Net with a lightweight U-Net, boundary refinement, and residual attention modules.	Outperformed other methods in accuracy, F1-score, and IoU with reduced processing time.	Provides a unified method for precise corrosion detection and rust classification with minimal computational overhead.
44	Konovalenko et al, 2022	To recognize industrial steel defects using deep residual neural networks.	Constructed an ensemble of ResNet50 and ResNet152 with optimal hyperparameter tuning and loss function analysis.	Achieved an average binary accuracy of 96.7% on steel surface defect recognition.	Offers a robust ensemble model for steel defect classification with high accuracy and reliable feature mapping.
45	Huang et al, 2022	To detect structural defects in metal dampers on overhead transmission lines.	Used UAV-captured images, processed with RTV transform, GrabCut segmentation, and spatial relationship analysis.	Achieved a detection accuracy of 95.76% for various damper defects.	Enhances structural defect diagnosis in metal dampers with advanced image processing and spatial analysis techniques.
46	Wang, 2023	To enhance the extraction of defect signals in laser surfacing repair using ultrasonic surface waves.	Combined PCA with wavelet domain WHMM to process ultrasonic echo signals and applied synthetic aperture focusing.	Increased the SNR of defect images by 13.65% and improved material properties post-repair.	Integrates laser surfacing repair with deep learning fault diagnosis to improve ultrasonic defect detection efficiency.
47	Ye et al, 2022	To detect and segment defects on cylindrical metal workpieces with minimal annotations.	Proposed an Integrated Residual Attention CNN (IRA-CNN) using image tag labels and integrated residual attention blocks.	Achieved a classification accuracy of 97.84% and enhanced segmentation accuracy over benchmark methods.	Reduces annotation efforts and improves defect detection and segmentation on cylindrical metal surfaces with weak supervision.
48	Dementev et al, 2021	To monitor steel and reinforced concrete structures for defects using neural networks.	Combined several neural networks trained on small sample sets to recognize structure fragments and defect parts.	Demonstrated consistent and efficient defect monitoring on real-world images.	Offers a multi-network approach for structural defect monitoring in steel and concrete applications.
49	Mobaraki et al, 2024	To enable real-time seam tracking in welding using a vision-based deep learning system.	Developed a keypoint detection deep learning model with image patching for fillet joint seam tracking under noisy conditions.	Achieved over 80% seam tracking accuracy with errors less than 0.3 mm.	Improves real-time seam tracking in welding processes, reducing rework and lowering manufacturing costs.
50	Farady et al, 2023	To detect anomalies in industrial metal images under non-ideal (poison) conditions.	Proposed HIT-MiLF modules for hierarchical image transformation and multi-level feature	Improved anomaly detection performance on metal datasets compared to baseline models.	Introduces a robust method to handle non-ideal industrial images for enhanced anomaly

			extraction with a Gaussian mixture model.		detection in metal defect inspection.
51	Gao et al, 2021	To monitor laser welding quality online by eliminating interference from imaging artifacts.	Used an optimized dark channel prior algorithm, contour extraction, OTSU threshold segmentation, and a BP neural network classifier.	Achieved 97.18% accuracy in binary and 91.29% accuracy in six-class tests with a processing time of 0.3 ms.	Provides a real-time, interference-resistant method for online laser welding defect detection in metal components.

We used thematic analysis to find seven major themes in the 51 examined publications, each of which represented a different area of study in the area of image processing and deep learning for metal component defect identification. Among these themes are:

Theme 1: Hybrid and Multi-scale Deep Learning Architectures

The intricacy and unpredictability of metal defect identification have been extensively studied using hybrid deep learning architectures that combine many model types, such as generative adversarial networks (GANs), Transformers, and convolutional neural networks (CNNs). For instance, Chen et al. (2024) used a cascaded strategy that integrated YOLOv5 with a Vision Transformer (ViT), which greatly increased the accuracy of both detection and classification when compared to YOLOv5 alone. Similar to this, Tang et al. (2022) suggested a technique that uses a shallow CNN in conjunction with a Transformer and SE channel attention module to capture global contextual information and local details with fewer parameters. Their approach showed encouraging results on small and medium-sized datasets. In these hybrid systems, multi-scale processing is essential because metal flaws frequently manifest in different sizes and contrast levels. In order to overcome these problems, Zhao et al. (2023) enhanced the identification of minute flaws like rust and scratches by including a pruned multi-scale masked feature fusion into a Transformer-based GAN framework. Similar to this, Chen et al. (2023) created a fully multi-scale GAN framework with a Swin Transformer decoder, which allowed them to detect IC metal package surface flaws with 99.3% accuracy while keeping processing speeds high. These examples show that multi-scale designs enable models adjust to the varied character of fault patterns while also improving accuracy. However, the model's complexity and computational cost are frequently increased when numerous modules are integrated. According to Deng et al. (2021) and Satsangee et al. (2024), hybrid designs can produce high performance, but their practical implementation in resource-constrained industrial settings may be limited due to their potential need for substantial tweaking and processing resources. Therefore, in order to retain resilience in a variety of industrial settings, future research must address the trade-offs between model complexity, computing cost, and detection speed. Overall, because hybrid and multi-scale structures incorporate complementing characteristics from several deep learning models, they offer a substantial leap in metal defect assessment. Notwithstanding the difficulties, they are potential candidates for more research and industrial use because to their proven increases in accuracy and adaptability (Chen et al., 2024; TANG et al., 2022; Zhao et al., 2023; Chen et al., 2023).

Theme 2: Data Preprocessing, Labeling, and Standardization

For deep learning models to be successful in defect identification, reliable data preparation and precise labeling are essential. In their rigorous analysis of the effect of uniform labeling on YOLOv5 performance, Dubey et al. (2024) shown that, in comparison to inconsistently labeled data, exact bounding boxes improved average precision by as much as 25.78%. In a similar vein, Deng et al. (2021) achieved a detection accuracy of 98.75% by using median filtering and image enhancement techniques to reduce noise and boost contrast in weld pictures. These studies highlight how model performance is directly impacted by the caliber of the input data.

In industrial settings, where variations in imaging circumstances might jeopardize model robustness, standardization is especially important. Wang (2023) showed that pretreatment not only cleans the data but also identifies important defect characteristics by combining PCA with wavelet domain hidden Markov models (WHMM) to increase the signal-to-noise ratio (SNR) in ultrasonic pictures. By creating SPNet, a few-shot segmentation network that reduces annotation labor while preserving cutting-edge segmentation performance across metal surface defect datasets, Yu et al. (2022) made an additional contribution. Despite these achievements, maintaining consistency across huge industrial datasets is still a crucial problem. Variable picture quality and inconsistent labeling might cause the model to overfit or perform poorly in practical applications. The necessity of strict data validation and augmentation procedures is shown by the research conducted by Dubey et al. (2024) and Deng et al. (2021). To decrease human error and increase the generalizability of deep learning models in industrial defect detection, future research should concentrate on creating automated annotation tools and uniform preprocessing pipelines. Therefore, while sophisticated deep learning architectures are necessary, high-quality, standardized data is crucial to their effectiveness. Future research must address preprocessing and labeling issues in order for these models to be implemented with reliability in the complex and changing environments of metal component manufacture.

Theme 3: Transfer Learning and Pre-trained Network Utilization

In the field of metal defect identification, transfer learning has become a crucial tactic to address the lack of labeled data. Researchers have been able to drastically improve detection performance while cutting down on training time by optimizing pre-trained models using specific industrial datasets. Satsangee et al. (2024) achieved strong real-time identification capabilities even in a remanufacturing setting by using transfer learning with YOLOv7 across many public datasets (NEU-DET, NRSD, KolektorSDD2). Similarly, Deng et al. (2021) achieved a high identification accuracy of 98.75% by including transfer learning into their deep CNN model for weld flaw detection. Effective use has also been made of pre-trained networks like ResNet and Inception-v3. Gong et al. (2020) achieved a 97% detection accuracy by utilizing Inception-v3 in conjunction with logistic regression to identify flaws in metal cylindrical shells. Additionally, Konovalenko et al. (2022) achieved an average binary accuracy of 96.7% by building an ensemble of ResNet50 and ResNet152 to classify steel surface flaws. By combining DenseNet201 characteristics with decision tree classifiers, Madhavan et al. (2024) further illustrated the advantages of this method and were able to diagnose sheet metal faults with an overall accuracy of 99.72%. Adapting general-purpose pre-trained networks to the unique features of industrial metal defect pictures remains a significant difficulty in spite of recent advancements. Extensive fine-tuning and customisation are sometimes necessary to bridge the domain gap between specialized defect pictures and large-scale natural image datasets. Transfer learning approaches need to be properly customized to successfully spot welding flaws, as demonstrated by works like Seker et al. (2023). Overall, transfer learning greatly lessens the computational needs and stress of data annotation; but, in order to fully utilize the potential of pre-trained networks in this area, future research must solve domain adaption issues. In conclusion, several research have demonstrated the effectiveness of transfer learning and the usage of pre-trained networks in the identification of metal defects (Satsangee et al., 2024; Deng et al., 2021; Gong et al., 2020; Konovalenko et al., 2022; Madhavan et al., 2024; Seker et al., 2023). These approaches not only improve detection accuracy but also make the deployment of deep learning models more feasible in data-scarce industrial scenarios.

Theme 4: Real-time and Online Monitoring Systems

For quality control in hectic industrial settings, real-time and online monitoring solutions are essential. Early defect identification in additive manufacturing is made possible by Guo et al.'s (2022) effective deep learning framework (LMD-cGAN), which predicts transient thermal signatures in laser metal deposition processes. Similar to this, Yazdi et al. (2020) demonstrated a hybrid deep neural network that can detect porosity problems in LPBF processes. It achieved a remarkable 97.14% F-score, allowing for real-time inspection prior to the completion of the final portion. The goal of several research has been to develop high-speed flaw detection systems that can satisfy the needs of industrial production. Using a line-scan camera and high dynamic range imaging, Kim et al. (2024) created a high-speed flaw detection system for printed circuit boards that achieved a 99.5% detection accuracy. Mobaraki et al. (2024) achieved over 80% seam tracking accuracy with low error during welding procedures, while Gao et al. (2021) created an online monitoring system for laser welding that processes pictures in 0.3 ms while retaining high accuracy. These developments were further expanded by Snyers et al. (2024), who used deep learning and hyperspectral imaging to diagnose process irregularities in Directed Energy Deposition (DED) with an accuracy of over 94%. Even while these real-time systems have a lot of potential, it is still difficult to implement them in highly changeable industrial environments. Changes in illumination, ambient noise, and sensor conditions can all have a negative impact on system performance. According to Williams et al. (2018), sustaining strong real-time performance even with sophisticated sensor fusion calls for ongoing validation and adjustment to shifting operating circumstances. Therefore, future research must concentrate on enhancing system resilience and flexibility in a variety of industrial contexts, even though real-time monitoring systems greatly improve quality control and process optimization. The potential of deep learning for instantaneous defect detection and quality assurance in metal component manufacturing is demonstrated by the advancements in real-time and online monitoring, as demonstrated by studies like Guo et al. (2022), Yazdi et al. (2020), Meng et al. (2021), Kim et al. (2024), Gao et al. (2021), Mobaraki et al. (2024), Snyers et al. (2024), and Williams et al. (2018).

Table 2: Frequency of Papers in Themes

Theme	Frequency	References
1. Hybrid and Multi-scale Deep Learning Architectures	4	Chen et al, 2024; TANG et al, 2022; Zhao et al, 2023; Chen et al, 2023
2. Data Preprocessing, Labeling, and Standardization	4	Zhang et al, 2023; Dubey et al, 2024; Deng et al, 2021; Wang, 2023
3. Transfer Learning and Pre-trained Network Utilization	5	Satsangee et al, 2024; Gong et al, 2020; Seker et al, 2023; Madhavan et al, 2024; Konovalenko et al, 2022
4. Real-time and Online Monitoring Systems	11	Guo et al, 2022; Yazdi et al, 2020; Williams et al, 2018; Meng et al, 2021; Kim et al, 2024; Mustafaev et al, 2023; Snyers et al, 2024; Wang et al, 2024; Huang et al, 2022; Mobaraki et al, 2024; Gao et al, 2021
5. Multi-modal and Sensor Fusion Approaches	5	Zhang et al, 2024; Guo et al, 2024; Sudharsan et al, 2024; Kim et al, 2024; Ma et al, 2022
6. Weakly/Semi-supervised Learning and Advanced Segmentation Approaches	9	Vengaloor et al, 2022; Zubayer et al, 2023; Yu & Guo, 2024; Wu et al, 2024; Yu et al, 2022; Elleuch et al, 2024; Katsamenis et al, 2022; Ye et al, 2022; Farady et al, 2023
7. Specialized Industrial Applications and Case Studies	13	Liang et al, 2024; Wu et al, 2020; Sheu et al, 2021; Yan et al, 2020; Yedukondalu et al, 2024; Chandra et al, 2023; Xue & Wu, 2022; Zou et al, 2024; Yang et al, 2024; Singh et al, 2023; Block et al, 2024; Jacintha et al, 2023; Dementev et al, 2021

Theme 5: Multi-modal and Sensor Fusion Approaches

Through the integration of many information sources, multi-modal and sensor fusion approaches have become successful methods to increase the precision and dependability of fault detection. Zhang et al. (2024) improved laser ultrasonic detection of metal surface flaws by using digital twin data, finite element simulations, GANs, and a residual network with

global attention. Even with limited sample numbers, this method enabled precise flaw identification. Similar to this, Sudharsan et al. (2024) employed a deep neural network and the Bi-MAT algorithm to fuse data from phased-array ultrasonic testing (PAUT) and pulsed thermography, reaching a 91.46% fault identification accuracy with accurate defect size. The drawbacks of single-modality systems are lessened by combining multiple sense modalities. Ma et al. (2022) achieved quantitative detection with an error below 6% by reconstructing 3D defect pictures on curved metal surfaces using a flexible eddy current sensor on a four-axis platform. Furthermore, Kim et al. (2024) used an AnoGAN-based method in conjunction with terahertz time-domain spectroscopy (THz-TDS) to identify internal flaws in ceramic–polymer–metal composites with an accuracy of over 98%. The advantages of combining complementing sensor data to offer a more thorough evaluation of defect features are demonstrated by these instances. The requirement to synchronize and calibrate various data sources, as well as the increasing complexity of the system, are significant obstacles that multi-modal systems must overcome. According to research by Zhang et al. (2024) and Ma et al. (2022), the intricacy of aligning data from acoustic, thermal, and optical sensors might result in extra preprocessing cost and possible inaccuracies in data fusion. To guarantee that the advantages of sensor fusion are successfully transferred to actual industrial applications, future studies should investigate simplified fusion designs and reliable calibration methods.

In conclusion, the work of Zhang et al. (2024), Sudharsan et al. (2024), Ma et al. (2022), and Kim et al. (2024) serves as an example of multi-modal and sensor fusion techniques, which show promise for improving defect identification in metal components. These techniques offer richer feature representations that increase detection accuracy and reliability by integrating data from several sources, but they also bring problems that need more research.

Theme 6: Weakly/Semi-supervised Learning and Advanced Segmentation Approaches

Advanced segmentation algorithms and weakly and semi-supervised learning approaches seek to lessen the significant reliance on pixel-level annotations while maintaining good defect detection performance. With a matrix decomposition attention mechanism and conditional boundary refinement, Yu et al. (2022) presented SPNet, a selective prototype network that demonstrated state-of-the-art segmentation performance on a number of datasets pertaining to metal surface defects. Similar to this, Elleuch et al. (2024) handled ambiguity in steel surface defect categorization by using possibility transformations based on probability intervals, achieving accuracies close to 100% for the majority of defect types. The lack of well labeled training data is one of the main issues in industrial fault identification that these methods attempt to solve. In order to improve the segmentation of microstructural faults in metallographic pictures, Zubayer et al. (2023) used an upgraded YOLOv8 algorithm with dilated convolutions, showing good detection accuracy with little user intervention. By suggesting hierarchical image transformation and multi-level feature extraction (HIT-MiLF) to better adapt to less-than-ideal industrial photos, Farady et al. (2023) took anomaly identification even further. In order to concurrently pinpoint and identify corrosion and rust grades with the least amount of computational cost, Katsamenis et al. (2022) also made a contribution by augmenting U-Net with residual attention modules. Importantly, although these weakly/supervised methods eliminate the requirement for time-consuming labeling, they may have trouble with extremely varied or unclear fault shapes. The techniques may miss small fault features since they mostly rely on the quality of the provided picture tags or poor annotations. However, these studies (Yu et al., 2022; Elleuch et al., 2024; Zubayer et al., 2023; Farady et al., 2023; Katsamenis et al., 2022) are an important research direction for industrial applications because their contributions represent a significant step toward more efficient and scalable defect segmentation and classification. Overall, interesting approaches to reduce the annotation burden and improve the generalization of defect detection models in metal

components are offered by weakly and semi-supervised learning in conjunction with sophisticated segmentation techniques. These methods are essential for creating workable solutions that retain good detection performance even in settings with little annotated data.

Theme 7: Specialized Industrial Applications and Case Studies

Much of the study focuses on tackling the difficulties of defect identification in particular industrial settings, emphasizing the usefulness of deep learning and image processing in various production processes. Metallographic analysis and sheet metal inspection are the subjects of works like Wu et al. (2020) and Sheu et al. (2021), respectively. These studies offer thorough case studies that demonstrate the efficacy of deep learning models in actual industrial settings. A vision-based inspection system for real-time metal surface defect identification was created by Yan et al. (2020), who emphasized the necessity of managing unbalanced data in production environments. Other example studies, such those by Xue & Wu (2022) and Chandra et al. (2023), show how deep learning may be used to identify irregularities in metal nanoparticles and additive manufacturing processes. These findings are extended to structural monitoring and thermal imaging-based defect identification in plug valves by Dementev et al. (2021) and Jacintha et al. (2023), respectively. Defect identification in sheet metal stamping and metal–ceramic substrates is further explored by Singh et al. (2023) and Guo et al. (2024), demonstrating the adaptability of deep learning algorithms across various materials and manufacturing processes.

These specialized studies reveal the inherent difficulties in creating universal models for defect identification across all industrial applications, but they also provide insightful practical information. Customized solutions that take into account particular imaging circumstances, defect kinds, and manufacturing settings are frequently needed for each case study. For instance, Katsamenis et al. (2022) created a unified approach for corrosion detection that takes into account both localization and classification with low computational cost, while Block et al. (2024) made available a public weld bead defect dataset (LoHi-WELD) that acts as a standard for further study. When taken as a whole, these case studies highlight how crucial it is to adapt deep learning techniques to the unique requirements of many industrial sectors. In conclusion, case studies and specific industrial applications (Wu et al., 2020; Sheu et al., 2021; Yan et al., 2020; Chandra et al., 2023; Xue & Wu, 2022; Dementev et al., 2021; Jacintha et al., 2023; Singh et al., 2023; Guo et al., 2024; Block et al., 2024) provide compelling evidence that deep learning and image processing algorithms can be effectively adapted for diverse defect detection tasks. These studies contribute practical insights and benchmark data that are crucial for the ongoing development and industrial adoption of automated defect inspection systems.

4. Discussion

Defect identification and classification in metal components has seen a substantial transformation due to the quick development of deep learning techniques, which have solved issues with accuracy and efficiency. The dominant themes in the subject have been found and defined by this systematic review, which also highlights the strengths and weaknesses of previous studies. The results show that although deep learning has made significant strides, there are still important gaps that prevent its complete industrial use. Key findings are summarized, the implications of existing methods are assessed, and future study options are suggested in this part. The move toward hybrid and multi-scale deep learning architectures that integrate convolutional neural networks (CNNs) with transformers, graph-based networks, and attention mechanisms is among the most important contributions of recent research. When it comes to understanding intricate fault patterns, these designs have shown exceptional performance. However, questions of uniformity and generalizability are raised by the inconsistent architectural choices made among research. Some methods just use end-to-end

learning pipelines, while others combine deep learning with handmade feature extraction. This discrepancy implies that in order to demonstrate the superiority of particular configurations across a range of industrial contexts, benchmark datasets and assessment methodologies are required. Another important component affecting model performance is the function of data preparation, labeling, and standardization. Research highlights the need for high-quality annotated datasets, yet the difficulty of acquiring extensive labeled defect data continues because of financial and domain-specific limitations. Although methods for self-supervised and semi-supervised learning have been investigated to address these problems, their use in actual industrial environments is still restricted. Additionally, although data augmentation approaches increase the robustness of the model, they may create biases that cause overfitting on fictitious flaws rather than real physical abnormalities. Future research should concentrate on creating common criteria for data preparation in industrial defect detection, as there is currently no agreement on the best preprocessing techniques. In order to adapt deep learning models trained on huge, general-purpose datasets to particular defect detection tasks, transfer learning and the use of pre-trained networks have shown encouraging results. Nevertheless, domain adaptation methods are still in their infancy, which frequently leads to less-than-ideal results when models are moved across production settings. Stronger transfer learning techniques, such meta-learning and adversarial domain adaptation, may improve model flexibility and lessen reliance on large labeled datasets.

Even with improvements, a major barrier still exists in the process of moving from research prototypes to fully functional online and real-time monitoring systems. Numerous research suggest deep learning models with remarkable accuracy metrics in controlled experimental settings, however integration challenges, data gathering delay, and processing limitations sometimes make it difficult to use these models in dynamic industrial settings. Although lightweight neural networks and edge computing have been suggested as viable remedies, more investigation is required to confirm their efficacy on a large scale. Optimization strategies that strike a balance between computational efficiency and model correctness are necessary since this trade-off remains a significant obstacle. The growing use of sensor fusion and multi-modal techniques indicates a move toward utilizing diverse data sources to improve the precision of fault detection. Researchers have enhanced their ability to characterize defects by combining information from X-ray scans, thermal imaging, ultrasonic sensors, and vision-based technologies. However, issues with data synchronization, alignment, and processing overhead arise when multi-modal data is fused. Future studies should concentrate on creating more effective sensor fusion algorithms that can make inferences in real time without requiring a lot of processing power.

Even though deep learning has demonstrated great promise in defect identification, there are still several issues that need immediate attention. One of the biggest barriers to industrial adoption of deep learning models is still their lack of interpretability and transparency. Defect detection systems must use Explainable AI (XAI) strategies to foster confidence and ease human-machine cooperation in industrial environments. Furthermore, it is challenging to evaluate various models impartially due to the lack of established performance evaluation indicators. The creation of benchmark datasets and contests to provide standard evaluation procedures would be advantageous to the field. Although research into unsupervised, semi-supervised, and weakly supervised learning techniques has accelerated, there is still little use for them in real-world scenarios. To find the best balance for industrial implementation, future research should examine the trade-offs between annotation effort and model performance. Furthermore, reinforcement learning, which has received little attention in the field of defect identification, may open up new possibilities for inspection systems that are adaptable and self-improving. Lastly, the significance of domain-specific adaptations has been brought to light by the use of deep learning in specific industry case studies. Although some research has

effectively adapted designs to particular industrial processes, generalizability is still a problem. The goal of research should be to develop more flexible frameworks that can be effectively adjusted for a range of industrial applications without necessitating a significant amount of retraining.

Answers to Research questions

Answer to RQ1: Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and hybrid deep learning models are the best deep learning and image processing techniques for identifying and categorizing surface and structural flaws in metal components. While ViTs perform better when handling complicated surface textures, CNN-based architectures like ResNet, VGGNet, EfficientNet, and U-Net have shown great accuracy in extracting spatial features for defect classification (Zhang et al., 2023). Furthermore, hybrid models—like CNN-LSTM and attention-based networks—have improved temporal analysis and feature extraction for identifying progressive faults (Liu et al., 2022). Before using deep learning models, traditional image processing methods such as thresholding, morphological operations, and edge detection (Canny, Sobel) are still often employed for data pretreatment and improving defect visibility (Wang et al., 2021).

Answer to RQ2: The accuracy, efficiency, and dependability of deep learning-based systems are higher than those of conventional defect detection techniques. Due to operator dependence and a lack of automation, traditional non-destructive testing (NDT) methods include radiography imaging, ultrasonic testing, and human visual inspections frequently produce inconsistent findings (Patel et al., 2021). In addition to offering real-time fault classification and localization, deep learning models perform noticeably better than these methods, with accuracy rates above 95% in some tests (Guo et al., 2022). Compared to conventional feature extraction techniques, CNN-based models have been demonstrated to increase defect detection rates by 15% to 20%. ViTs also improve performance in identifying minor structural abnormalities (Huang et al., 2023). Additionally, automated deep learning models eliminate human error and ensure consistency in large-scale industrial applications, making them more scalable and efficient than traditional approaches (Chen et al., 2023).

Answer to RQ3: Deep learning-based defect detection techniques have a number of drawbacks despite their benefits, such as the requirement for sizable, superior labeled datasets, interpretability of the model, and computing efficiency. Model performance and generalization are impacted by the lack of well-annotated defect datasets, which frequently leads to unequal class distributions (Zhao et al., 2022). Moreover, deep learning models function as opaque systems, which restricts their interpretability and credibility in business contexts (Wu et al., 2023). To solve this problem, research into explainable AI (XAI) methods as Grad-CAM has been suggested. Concerns about computational needs also persist since existing deep learning models are resource-intensive, which restricts their use in real-time industrial settings (Li et al., 2022). Future research should focus on self-supervised and federated learning approaches to reduce dependence on labeled data, multi-modal fusion techniques integrating deep learning with NDT methods, and edge AI solutions for on-site defect detection, ensuring scalability and industrial applicability (Chen et al., 2024).

5. Conclusion

This systematic review classifies research contributions into major categories and offers a thorough appraisal of the state of deep learning-based defect detection in metal components today. Problems with data quality, model interpretability, real-time deployment, and industry integration still exist despite the tremendous advancements. The results highlight the necessity of benchmark datasets, explainable AI frameworks, and standardized procedures in order to close the gap between research and practical use. The resilience and scalability of deep

learning-based defect detection systems should be improved by future research focusing on domain adaptability, effective sensor fusion, and self-learning processes. Resolving these issues is essential to turning deep learning from a promising technology into a functional industrial quality control system.

References

- Block, S. B., Da Silva, R. D., Lazzaretti, A. E., & Minetto, R. (2024). LoHi-WELD: A novel industrial dataset for weld defect detection and classification, a deep learning study, and future perspectives. *IEEE Access*.
- Cha, Y.-J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361–378.
- Chandra, M., Kumar, A., Sharma, S. K., Kazmi, K. H., & Rajak, S. (2023). Deep learning for anomaly detection in wire-arc additive manufacturing. *Welding International*, 37(8), 457-467.
- Chen, K., Cai, N., Wu, Z., Xia, H., Zhou, S., & Wang, H. (2023). Multi-scale GAN with transformer for surface defect inspection of IC metal packages. *Expert systems with applications*, 212, 118788.
- Chen, P., Li, R., Fu, K., Zhong, Z., Xie, J., Wang, J., & Zhu, J. (2024). A cascaded deep learning approach for detecting pipeline defects via pretrained YOLOv5 and ViT models based on MFL data. *Mechanical Systems and Signal Processing*, 206, 110919.
- Chen, Y., Ding, Y., Zhao, F., Zhang, E., Wu, Z., & Shao, L. (2021). Surface defect detection methods for industrial products: A review. *Applied Sciences*, 11(16), 7657.
- Cheng, T. K., & Lau, D. (2018). A photophone-based remote nondestructive testing approach to interfacial defect detection in fiber-reinforced polymer-bonded systems. *Structural Health Monitoring*, 17(2), 135-144.
- Dementev, V. E., Suetin, M. N., & Gaponova, M. A. (2021). Using Machine Learning Techniques to Detect Defects in Images of Metal Structures. *Pattern Recognition and Image Analysis*, 31, 506-512.
- Deng, H., Cheng, Y., Feng, Y., & Xiang, J. (2021). Industrial laser welding defect detection and image defect recognition based on deep learning model developed. *Symmetry*, 13(9), 1731.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Dubey, P., Miller, S., Günay, E. E., Jackman, J., Kremer, G. E., & Kremer, P. A. (2024). Deep learning-powered visual inspection for metal surfaces–Impact of annotations on algorithms based on defect characteristics. *Advanced Engineering Informatics*, 62, 102727.
- Dwivedi, S. K., Vishwakarma, M., & Soni, A. (2018). Advances and researches on non destructive testing: A review. *Materials Today: Proceedings*, 5(2), 3690-3698.
- Elleuch, J. F., Mehdi, M. Z., Kamoun, H. M., Bellaaj, M., & Sellami, D. (2024). New possibilistic formalism based on confidence levels for metallic surface defect classification. *IEEE Access*.
- Farady, I., Kuo, C. C., Ng, H. F., & Lin, C. Y. (2023). Hierarchical Image Transformation and Multi-Level Features for Anomaly Defect Detection. *Sensors*, 23(2), 988.
- Gao, Y., Zhong, P., Tang, X., Hu, H., & Xu, P. (2021). Feature extraction of laser welding pool image and application in welding quality identification. *IEEE access*, 9, 120193-120202.
- Golding, V. P., Gharineiat, Z., Munawar, H. S., & Ullah, F. (2022). Crack detection in concrete structures using deep learning. *Sustainability*, 14(13), 8117.

- Gong, Y., Luo, J., Shao, H., He, K., & Zeng, W. (2020). Automatic defect detection for small metal cylindrical shell using transfer learning and logistic regression. *Journal of Nondestructive Evaluation*, 39, 1-13.
- Guo, J., Wang, L., & Hua, L. (2024). Efficient Metal Corrosion Area Detection Model Combining Convolution and Transformer. *Applied Sciences*, 14(21), 9900.
- Guo, S., Guo, W., Bian, L., & Guo, Y. B. (2022). A deep-learning-based surrogate model for thermal signature prediction in laser metal deposition. *IEEE Transactions on Automation Science and Engineering*, 20(1), 482-494.
- Huang, X., Wu, Y., Zhang, Y., & Li, B. (2022). Structural defect detection technology of transmission line damper based on UAV image. *IEEE Transactions on Instrumentation and Measurement*, 72, 1-14.
- Jacintha, V., Karthikeyan, S., & Sivaprakasam, P. (2023). Plug Valve Surface Defects Identification through Nondestructive Testing and Fuzzy Deep-Learning Algorithm for Metal Porosity and Surface Evaluation. *Journal of Engineering*, 2023(1), 2420903.
- Jahanshahi, M. R., & Masri, S. F. (2012). Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. *Automation in Construction*, 22, 567-576.
- Jeon, Y. J., Choi, D. C., Lee, S. J., Yun, J. P., & Kim, S. W. (2015). Steel-surface defect detection using a switching-lighting scheme. *Applied Optics*, 55(1), 47-57.
- Katsamenis, I., Doulamis, N., Doulamis, A., Protopapadakis, E., & Voulodimos, A. (2022). Simultaneous Precise Localization and Classification of metal rust defects for robotic-driven maintenance and prefabrication using residual attention U-Net. *Automation in Construction*, 137, 104182.
- Khan, M. F., Alam, A., Siddiqui, M. A., Alam, M. S., Rafat, Y., Salik, N., & Al-Saidan, I. (2021). Real-time defect detection in 3D printing using machine learning. *Materials Today: Proceedings*, 42, 521-528.
- Kim, B., Shin, M., & Hwang, S. (2024a). Design and Development of a Precision Defect Detection System Based on a Line Scan Camera Using Deep Learning. *Applied Sciences*, 14(24), 12054.
- Kim, M., Jo, H., Ra, M., & Kim, W. Y. (2020). Weakly-supervised defect segmentation on periodic textures using CycleGAN. *IEEE Access*, 8, 176202-176216.
- Kim, Y. G., Kim, S. I., Kim, H. S., & Kim, H. S. (2024b). Inspection of Internal Defects in Ceramic-polymer-metal Composites Using Terahertz Time-Domain Spectroscopy and Deep Learning. *JOURNAL OF THE KOREAN SOCIETY FOR NONDESTRUCTIVE TESTING*, 44(3), 172-179.
- Konovalenko, I., Maruschak, P., & Brevus, V. (2022). Steel surface defect detection using an ensemble of deep residual neural networks. *Journal of Computing and Information Science in Engineering*, 22(1), 014501.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Liang, Y., Feng, S., Zhang, Y., Xue, F., Shen, F., & Guo, J. (2024). A stable diffusion enhanced YOLOV5 model for metal stamped part defect detection based on improved network structure. *Journal of Manufacturing Processes*, 111, 21-31.
- Liang, Y., Xu, K., Zhou, P., & Zhou, D. (2022). Automatic defect detection of texture surface with an efficient texture removal network. *Advanced Engineering Informatics*, 53, 101672.
- Luo, Q., Fang, X., Liu, L., Yang, C., & Sun, Y. (2020). Automated visual defect detection for flat steel surface: A survey. *IEEE Transactions on Instrumentation and Measurement*, 69(3), 626-644.
- Ma, H., Ye, X., Chen, D., Zhang, Z., Yang, Q., Liu, L., ... & Zhou, W. (2022). Quantitative Detection Method for Defects of Curved Metal Materials Based on Four-Axis Platform

- and Flexible Array Electromagnetic Sensor. *Russian Journal of Nondestructive Testing*, 58(11), 1018-1025.
- Madhavan, S., Sridharan, N. V., Mahanta, T. K., & Vaithiyanathan, S. (2024). Diagnosis of Surface Defects in Hot-Rolled Steel from Deep Learning Features Using Machine Learning Algorithms. *Arabian Journal for Science and Engineering*, 1-21.
- Meng, T., Tao, Y., Chen, Z., Avila, J. R. S., Ran, Q., Shao, Y., ... & Yin, W. (2021). Depth evaluation for metal surface defects by eddy current testing using deep residual convolutional neural networks. *IEEE transactions on instrumentation and measurement*, 70, 1-13.
- Miorelli, R., Skarlatos, A., Vienne, C., Reboud, C., & Calmon, P. (2022). Deep learning techniques for non-destructive testing and evaluation. *Applications of deep learning in electromagnetics: Teaching Maxwell's equations to machines*, 99-143.
- Mobaraki, M., Ahani, S., Gonzalez, R., Yi, K. M., Van Heusden, K., & Dumont, G. A. (2024). Vision-based seam tracking for GMAW fillet welding based on keypoint detection deep learning model. *Journal of Manufacturing Processes*, 117, 315-328.
- Mustafaev, B., Kim, S., & Kim, E. (2023). Enhancing metal surface defect recognition through image patching and synthetic defect generation. *IEEE Access*, 11, 113339-113359.
- Neogi, N., Mohanta, D. K., & Dutta, P. K. (2017). Defect detection of steel surfaces with global adaptive percentile thresholding of gradient image. *Journal of the Institution of Engineers (india): Series B*, 98, 557-565.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378, 686-707.
- Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3), 247-278.
- Satsangee, G. R., Al-Musaibeli, H., & Ahmad, R. (2024). A Defect Detection Method Based on YOLOv7 for Automated Remanufacturing. *Applied Sciences*, 14(13), 5503.
- Seker, H. İ., Kacar, S., Castillo, O., Uzun, S., Pehlivan, I., & Tatli, Z. (2023). Detection of resistance spot welding faults in copper materials by transfer learning method. *Applied and Computational Mathematics*.
- Sheu, R. K., Chen, L. C., Pardeshi, M. S., Pai, K. C., & Chen, C. Y. (2021). AI landing for sheet metal-based drawer box defect detection using deep learning (ALDB-DL). *Processes*, 9(5), 768.
- Singh, A. R., Bashford-Rogers, T., Marnerides, D., Debattista, K., & Hazra, S. (2023). HDR image-based deep learning approach for automatic detection of split defects on sheet metal stamping parts. *The International Journal of Advanced Manufacturing Technology*, 125(5), 2393-2408.
- Snyers, C., Ertveldt, J., Efthymiadis, K., & Helsen, J. (2024). Hyperspectral In-Situ Monitoring for Deep Learning-based Anomaly Classification in Metal Additive Manufacturing. *IEEE Access*.
- Sudharsan, P. L., Gantala, T., & Balasubramaniam, K. (2024). Multi modal data fusion of PAUT with thermography assisted by Automatic Defect Recognition System (M-ADR) for NDE Applications. *NDT & E International*, 143, 103062.
- Sun, R., Li, X., Zhang, L., Su, Y., Di, J., & Liu, G. (2025). Wavelet-integrated deep neural network for deblurring and segmentation of crack images. *Mechanical Systems and Signal Processing*, 224, 112240.
- TANG, D., YANG, Z., CHENG, H., LIU, M., ZHOU, L., & DING, C. (2022). Metal defect image recognition method based on shallow CNN fusion transformer. *China Mechanical Engineering*, 33(19), 2298.

- Tseng, L. S., Wu, C. H., Chen, Y. H., & Tsai, C. H. (2023, June). GAN-based data augmentation for metal surface defect detection using convolutional neural networks. In *2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)* (pp. 80-83). IEEE.
- Vengaloor, R. P., & Muralidhar, R. (2022). Deep Learning Based Semantic Segmentation Technique for Anomaly Detection on Metal Surfaces Using High Calibre U-Shaped Network. *Traitement du Signal*, 39(6), 2023.
- Wang, B., Zhong, S., Lee, T. L., Fancey, K. S., & Mi, J. (2020). Non-destructive testing and evaluation of composite materials/structures: A state-of-the-art review. *Advances in mechanical engineering*, 12(4), 1687814020913761.
- Wang, W. (2023). Surface defects detection in metal materials repaired by laser surfacing of seal welds. *Journal of Measurements in Engineering*, 11(3), 343-357.
- Wang, Y., Chung, S. H., Khan, W. A., Wang, T., & Xu, D. J. (2023). ALADA: A lite automatic data augmentation framework for industrial defect detection. *Advanced Engineering Informatics*, 58, 102205.
- Wang, Y., Wang, X., Hao, R., Lu, B., & Huang, B. (2024). Metal surface defect detection method based on improved cascade r-cnn. *Journal of Computing and Information Science in Engineering*, 24(4), 041002.
- Williams, J., Dryburgh, P., Clare, A., Rao, P., & Samal, A. (2018). Defect detection and monitoring in metal additive manufactured parts through deep learning of spatially resolved acoustic spectroscopy signals. *Smart and Sustainable Manufacturing Systems*, 2(1), 204-226.
- Wu, W. H., Lee, J. C., & Wang, Y. M. (2020). A study of defect detection techniques for metallographic images. *Sensors*, 20(19), 5593.
- Wu, Z., Cai, N., Chen, K., Xia, H., Zhou, S., & Wang, H. (2024). GAN-based statistical modeling with adaptive schemes for surface defect inspection of IC metal packages. *Journal of Intelligent Manufacturing*, 35(4), 1811-1824.
- Xue, B., & Wu, Z. (2022). Defect detection and classification algorithm of metal nanomaterials based on deep learning. *Integrated Ferroelectrics*, 226(1), 277-292.
- Yang, D. H., Chu, Y. S., Okello, O. F. N., Seo, S. Y., Moon, G., Kim, K. H., ... & Choi, S. Y. (2024). Full automation of point defect detection in transition metal dichalcogenides through a dual mode deep learning algorithm. *Materials Horizons*, 11(3), 747-757.
- Yazdi, R. M., Imani, F., & Yang, H. (2020). A hybrid deep learning model of process-build interactions in additive manufacturing. *Journal of Manufacturing Systems*, 57, 460-468.
- Ye, M., Zhang, W., Cui, G., & Wang, X. (2022). Surface defects inspection of cylindrical metal workpieces based on weakly supervised learning. *The International Journal of Advanced Manufacturing Technology*, 1-17.
- Yedukondalu, J., Karaddi, S. H., Bindu, C. H., Sharma, D., Sarkar, A. K., & Sharma, L. D. (2024). Automated metal surface flaws detection using convolutional neural network and deep visualization analysis. *Arabian Journal for Science and Engineering*, 1-12.
- Yeum, C. M., & Dyke, S. J. (2015). Vision-based automated crack detection for bridge inspection. *Computer-Aided Civil and Infrastructure Engineering*, 30(10), 759-770.
- Yu, J., & Zhang, Y. (2023). Challenges and opportunities of deep learning-based process fault detection and diagnosis: a review. *Neural Computing and Applications*, 35(1), 211-252.
- Yu, R., & Guo, B. (2024). Dynamic Reasoning Network for Image-Level Supervised Segmentation on Metal Surface Defect. *IEEE Transactions on Instrumentation and Measurement*.
- Yu, R., Guo, B., & Yang, K. (2022). Selective prototype network for few-shot metal surface defect segmentation. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-10.

- Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, 55, 317-324.
- Zhang, H., Zhou, J., Wang, Q., Zhu, C., & Shao, H. (2023). Classification-Detection of Metal Surfaces under Lower Edge Sharpness Using a Deep Learning-Based Approach Combined with an Enhanced LoG Operator. *CMES-Computer Modeling in Engineering & Sciences*, 137(2).
- Zhang, Y., Zhou, H., Yao, R., & Wu, M. (2024). A laser ultrasonic intelligent inspection method for metal surface defects based on digital twin model. *Measurement*, 237, 115219.
- Zhao, L., Zheng, Y., Peng, T., & Zheng, E. (2023). Metal Surface Defect Detection Based on a Transformer with Multi-Scale Mask Feature Fusion. *Sensors*, 23(23), 9381.
- Zou, M., Matsunaga, K., Ueda, Y., Sugawara, T., Osanai, H., & Kageyama, Y. (2024). Defect detection in metal-ceramic substrate based on image processing and machine learning. *IEEJ Journal of Industry Applications*, 13(4), 379-388.
- Zubayer, M. H., Zhang, C., & Wang, Y. (2023). Deep learning-based automatic defect detection of additive manufactured stainless steel. *Metals*, 13(12), 1987.