IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 11 August 2024, accepted 29 August 2024, date of publication 2 September 2024, date of current version 10 September 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3453664

SURVEY

Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes

MD RAISUL ISLAM[1], MD ZAKIR HOSSAIN ZAMIL[2], MD ESHMAM RAYED[3], MD MOHSIN KABIR[4], M. F. MRIDHA[3], (Senior Member, IEEE), SATOSHI NISHIMURA[5], (Member, IEEE), AND JUNGPIL SHIN[5], (Senior Member, IEEE)

[1]School of Engineering and Technology, Western Illinois University, Macomb, IL 61455, USA
[2]School of Computer Sciences, Western Illinois University, Macomb, IL 61455, USA
[3]Department of Computer Science and Engineering, American International University-Bangladesh, Dhaka 1229, Bangladesh
[4]Faculty of Informatics, Eötvös Loránd University, 1117 Budapest, Hungary
[5]School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu 965-8580, Japan

Corresponding author: Jungpil Shin (jpshin@u-aizu.ac.jp)

ABSTRACT Ensuring product quality and integrity is paramount in the rapidly evolving landscape of industrial manufacturing. Although effective to a certain degree, traditional quality control methods often fail to meet the demands for efficiency, accuracy, and adaptability in today's fast-paced production environments. The advent of Deep Learning (DL) and Computer Vision (CV) technologies has opened new vistas for automated defect detection, promising to revolutionize the way industries approach quality control and inspection. This systematic review focuses on recent advancements in DL and CV applications for automated defect detection in manufacturing processes. It provides a comprehensive overview of state-of-the-art techniques for detecting, classifying, and predicting defects, highlighting the significant strides made in addressing challenges such as varying lighting conditions, complex defect patterns, and the seamless integration of these technologies into existing manufacturing workflows. Through a critical analysis of current methodologies, this study identifies key areas of opportunity, outlines the challenges that persist and suggests directions for future research. This review synthesizes findings from a broad spectrum of industrial applications, offering insights into the potential of DL and CV to enhance quality control mechanisms. By charting the progress and pinpointing the gaps in current practices, this paper aims to serve as a valuable resource for researchers, practitioners, and policymakers seeking to leverage the benefits of DL and CV for improved product management and manufacturing excellence.

INDEX TERMS Automated defect detection, industrial automation, manufacturing quality control, deep learning, computer vision, pattern recognition.

I. INTRODUCTION

Quality control remains a cornerstone in manufacturing, ensuring that products meet rigorous quality and reliability standards. Traditional quality control methods rely heavily on manual inspection and simple automated systems, which are time-consuming, prone to human error, and often lack the precision to detect complex defects [1]. The significance of quality control cannot be overstated, as it directly impacts product reliability, customer satisfaction, and the overall efficiency of manufacturing processes. The advent of Deep Learning (DL) and Computer Vision (CV) technologies has marked a paradigm shift in how industries approach defect detection and quality control [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

Quality control in manufacturing encompasses activities and procedures to ensure that products meet specified quality criteria and satisfy customer expectations. This is a critical component of the manufacturing process that influences product quality and efficiency. Traditional quality control methods, including human visual inspections and simple automated systems, have been integral to maintaining standards. However, these methods often fall short in today's complex manufacturing environments, where the diversity of products and the intricacies of defects require more sophisticated detection techniques. The limitations of traditional quality control methods, such as their labor-intensive nature, subjectivity, and inability to detect subtle or complex defects, underscore the need for advanced and reliable solutions.

Deep Learning and Computer Vision represent the forefront of technological advancements in industrial applications, offering unprecedented capabilities in automating and enhancing quality control processes [3], [4]. DL, a subset of machine learning, utilizes neural networks with multiple layers (deep networks) to learn from vast amounts of data, thereby enabling the automatic detection and classification of defects with high accuracy [5]. However, Computer Vision allows machines to interpret and understand visual information from the world, making it possible to identify, locate, and classify defects in manufacturing products by analyzing images and videos [6]. Integrating DL and CV into industrial applications has led to the development of sophisticated automated defect detection systems [7]. These systems can overcome many limitations associated with traditional quality control methods, offering advantages such as improved detection accuracy, learning from new data, and handling of complex defect patterns under varying conditions.

This systematic review explores the recent advancements in DL and CV technologies for automated defect detection in manufacturing processes. Table 1 compares our findings with those of existing surveys relevant to this field. The objectives are to:

- Provide a comprehensive overview of DL and CV techniques applied in defect detection and quality control, highlighting their strengths and limitations.
- Review state-of-the-art methodologies for detecting, classifying, and predicting manufacturing defects, focusing on challenges such as varying lighting conditions, complex defect patterns, and integration with existing manufacturing workflows.
- Identify gaps in current research and suggest directions for future studies to advance the automated defect detection domain.

Moreover, the scope of this review encompasses a broad range of DL and CV applications in various manufacturing sectors, including the automotive, electronics, and aerospace industries. Section II describes the methodology used in this systematic review. Section III examines various application domains of enhanced quality control in the manufacturing process. Section IV discusses the preprocessing techniques and algorithms used in the manufacturing process. Section V reviews recent advancements highlighted in the state-of-the-art literature. Section VI discusses the emerging trends and future directions in the field. Section VII identifies the ongoing challenges and limitations of the current study. Finally, Section VIII concludes the paper.

II. METHODOLOGY

This section outlines the systematic methodology employed in this review to collate, analyze, and synthesize the literature on advancements in automated defect detection using Deep Learning and Computer Vision techniques for quality control in manufacturing processes.

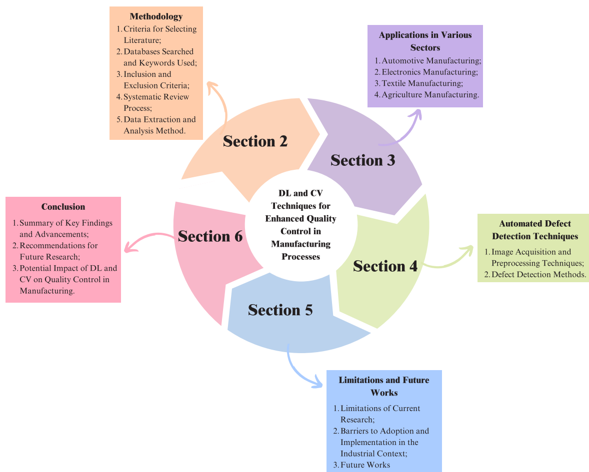

FIGURE 1. The structure of the paper.

A. CRITERIA FOR SELECTING LITERATURE

1) DATABASES SEARCHED

A literature search was carried out in several scientific databases and digital libraries, including IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar. These platforms were chosen for their comprehensive coverage of peer-reviewed articles, conference proceedings, and computer science, engineering, and industrial application journals.

2) KEYWORDS USED

A combination of keywords and phrases was used to ensure a comprehensive search. These included "automated defect detection," "deep learning in manufacturing," "computer vision quality control," "DL and CV in industrial applications," "defect classification using machine learning," and "automated inspection systems." Boolean operators (AND and, OR) were employed to refine the search results.

3) INCLUSION/EXCLUSION CRITERIA

The survey utilized essential resources obtained in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [13],

TABLE 1. Comparative performance and limitations of existing DL and CV based enhanced quality control in manufacturing process survey papers.

Ref	Applications	Pre-processing Techniques	Algorithms	Recent Works	Challenge & Future Direction	Limitation
Jha *et al.* [8]	✓	✗	✓	✗	✓	The study lacks a detailed discussion on specific pre-processing techniques, and although it includes result analysis, it does not cover recent state-of-the-art developments in the field.
Tercan *et al.* [9]	✓	✓	✓	✗	✓	The study does not provide a broad discussion on specific application areas within manufacturing, nor does it extensively cover pre-processing techniques crucial for data preparation. Furthermore, the paper lacks a broad analysis of DL & ML algorithms and does not include a recent state-of-the-art work analysis, which is essential for benchmarking and understanding current advancements in the field.
Yang *et al.* [10]	✓	✗	✓	✓	✓	The study does not detail the pre-processing techniques used to prepare the data for optimal model performance. Although it provides an analysis of different models, it lacks a thorough review of recent state-of-the-art works.
Nasir *et al.* [11]	✓	✗	✓	✗	✓	The study does not provide detailed information on the pre-processing techniques used for data preparation and lacks an in-depth analysis of recent state-of-the-art advancements.
Wang *et al.* [12]	✓	✗	✓	✗	✓	The study primarily include the lack of a broad discussion on the application areas and pre-processing techniques for deep learning in manufacturing. Furthermore, the study does not include any analysis of recent state-of-the-art works, which could have provided insights into current advancements and benchmarks in the field.
This Study	✓	✓	✓	✓	✓	-

TABLE 2. Criteria for selecting literature.

Inclusion Criteria	Exclusion Criteria
Papers published in the last five years	Papers not directly related to manufacturing processes
Studies addressing DL and CV in defect detection	Studies focusing on theoretical aspects without practical applications
Articles providing empirical evidence of effectiveness	Articles that were not peer-reviewed
Works available in English	Studies not available in English
Research focusing on a variety of manufacturing sectors	Papers with a narrow focus not applicable to broader manufacturing contexts
Studies including comparisons with traditional methods	Literature lacking comparative analysis with existing methods
Articles detailing integration strategies for DL and CV technologies into existing workflows	Papers overlooking the integration challenges of new technologies
Research presenting scalable DL and CV solutions	Studies not addressing scalability and real-world applicability
Articles with a clear methodology and result section	Works with vague methodologies or inconclusive results

as illustrated in figure 2. Specific inclusion and exclusion criteria governed the selection of the literature to ensure relevance and quality. The selection criteria are presented in Table 2 presents.

B. OVERVIEW OF THE SYSTEMATIC REVIEW PROCESS

The systematic review process followed predefined steps to ensure the comprehensiveness and reproducibility of the results. Initially, a broad search was conducted using the defined keywords across the selected databases. The search results were then screened based on titles and abstracts to exclude irrelevant papers. The remaining articles were subjected to a full-text review to assess suitability based on the inclusion and exclusion criteria. References of the selected articles were also scanned to identify additional relevant studies that might have been missed in the initial search.

C. DATA EXTRACTION AND ANALYSIS METHODS

1) DATA EXTRACTION

For each selected article, data were extracted on key aspects such as the authors, year of publication, study objectives, methodologies employed (specific DL and CV techniques), findings, and any noted challenges or limitations. This information was catalogued in a structured format to facilitate analysis.

2) ANALYSIS METHODS

The extracted data were analyzed using a qualitative approach. This study aimed to identify common themes, patterns, and trends in the selected literature. This included categorizing studies based on the types of manufacturing processes examined, specific DL and CV techniques applied, and outcomes achieved. The analysis also evaluated the challenges and limitations associated with these technologies in the context of industrial defect detection.

A critical review was conducted to assess the methodological quality of the studies, their applicability to different manufacturing sectors, and the robustness of the evidence provided. This comprehensive analysis allowed for the identification of gaps in the current literature and formulation of recommendations for future research in the field of automated defect detection using DL and CV.

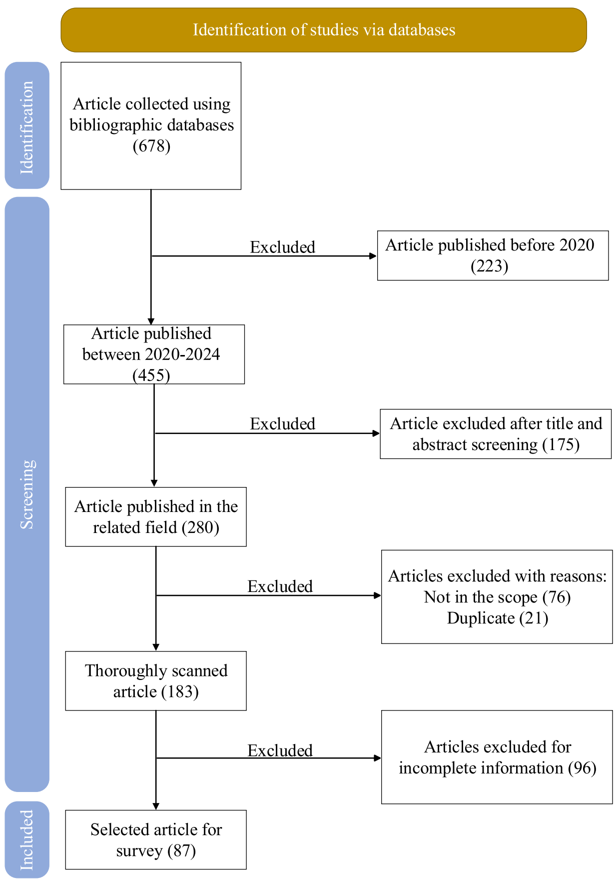

FIGURE 2. PRISMA flow diagram of paper selection process for the state-of-the-art works.

III. APPLICATIONS AND CASE STUDIES

Integrating Deep Learning and Computer Vision technologies has significantly enhanced quality control capabilities across various manufacturing sectors. This section explores their applications in the automotive, electronics, textiles, and agriculture sectors, showcasing the breadth of their impact and potential for further advancements.

A. AUTOMOTIVE MANUFACTURING

The automotive sector is at the forefront of adopting advanced technologies to improve quality control and manufacturing efficiency. Deep Learning and Computer Vision technologies have become pivotal in transforming traditional inspection processes, offering unprecedented accuracy and speed in detecting defects and ensuring the production of high-quality vehicles. This section discusses the specific applications, benefits, and advancements of DL and CV in automotive manufacturing, underlining their critical role in the evolution of the industry.

1) CASE STUDIES: AUTOMATED VISUAL INSPECTION SYSTEMS

Automated visual inspection systems powered by DL and CV, are designed to identify defects that are challenging for human inspectors to detect because of their small size or complex nature. These systems are employed at various stages of the automotive production process, from the inspection of incoming raw materials to the final assembly check. For instance, Wang and Gan [14] introduced a novel 2D-3D computer vision approach that leverages transfer learning for both the 3D reconstruction and visual inspection of buildings. This methodology demonstrates the versatility of DL and CV technologies beyond traditional manufacturing contexts. Specifically, in the automotive production process, DL algorithms play a crucial role in inspecting metal surfaces for imperfections, such as scratches or inconsistencies in paintwork, thereby ensuring that the final products adhere to the highest quality standards. Extending the application of these technologies, Yu et al. [15] presented an innovative corrosion detection system optimized for Micro Aerial Vehicles (MAVs). This system, named AMCD, underscores the adaptability of DL-based solutions to various inspection tasks and environments. From these examples, it is evident that automated visual inspection systems, powered by DL and CV, are instrumental in enhancing quality control measures across different stages of manufacturing, offering unparalleled precision and efficiency.

2) CASE STUDIES: DEFECT DETECTION IN ENGINE COMPONENTS

Engine components, which are crucial for the vehicle's performance and safety, undergo rigorous inspection. CV technologies facilitate the detection of casting defects, irregularities in machining, and assembly issues in components such as, pistons, cylinders, and crankshafts. To address the challenges of defect detection, Li et al. [16] introduced an enhanced YOLOv5 algorithm, DDSC-YOLOv5s, specifically designed for pinpointing defects in aero-engine blades and vanes. This approach represents a significant step forward in applying DL for the precise inspection of complex engine parts. Building on the integration of DL in inspection processes, Upadhyay et al. [17] developed an innovative automated framework for borescope inspections of aircraft engines. Their method combines deep learning with traditional computer vision techniques to offer a more comprehensive and accurate defect detection solution. Further contributing to the field, Wang et al. [18] proposed a novel detection framework, DBFF-YOLOv4, which leverages deep learning for radiographic testing of aeroengine turbine blades. This framework underscores the potential of DL models in enhancing the reliability of defect detection in critical engine components. DL models trained on thousands of images of defective and non-defective components can accurately identify potential failures, significantly reducing the risk of engine malfunctions. Building on this progress, Keshun et al. [19] proposed a novel hybrid model combining a Quadratic Neural Network (QNN) with a Bidirectional Long Short-Term Memory (Bi-LSTM) network. This model aims to improve the efficiency and interpretability of rolling bearing fault diagnosis in rotating machinery, leveraging QNN for multilayer feature extraction and Bi-LSTM for capturing the dynamic evolution of signals. The approach enhances fault diagnosis accuracy and speed while providing better interpretability through visualization techniques. In another study, Keshun et al. [20] proposed a deep learning-based probabilistic prediction model for estimating the remaining useful life (RUL) of machinery and equipment. The model aims to address the limitations of traditional RUL prediction methods by introducing a flexible prior distribution and a strategy for sequential optimization of hyperparameters. They used a modified ResNet architecture to improve the prediction performance, incorporating techniques to handle uncertainty and variability in real-world data. The model demonstrated superior prediction accuracy and robustness compared to traditional methods, specifically validated on the C-MAPSS dataset for turbofan engines.

3) CASE STUDIES: WELDING QUALITY ASSESSMENT

Welding is a fundamental process in automotive manufacturing that, requires precision and consistency. DL and CV systems were used to assess the quality of welding seams, identifying discontinuities, and predicting potential weak points in the structure of the vehicle. For instance, El Hachem et al. [21] introduced an innovative automated approach tailored for inspecting welding seams in the automotive sector, leveraging deep learning to enhance inspection accuracy and efficiency. Building on this, Miao et al. [22] developed a two-stage convolutional neural network (CNN) based method specifically designed for the online inspection of narrow overlap weld quality. This method exemplifies the application of DL in facilitating real-time evaluations of welding processes. Moreover, Wang et al. [23] proposed an advanced model based on an enhanced Faster R-CNN, integrating a Feature Pyramid Network (FPN), variable convolution, and a background suppression function to detect weld defects with higher precision. This approach underscores the potential of DL and CV in identifying intricate welding flaws that might compromise structural safety. Through these technological advancements, DL and CV systems enable continuous monitoring of the welding process, allowing for immediate adjustments and significantly reducing the likelihood of structural defects. This real-time feedback loop is instrumental in maintaining the high quality and safety standards required in automotive manufacturing.

B. ELECTRONICS MANUFACTURING

The electronics manufacturing industry, characterized by its intricate components and need for precision, has embraced Deep Learning and Computer Vision technologies to enhance quality control and efficiency [24], [25], [26]. This section outlines the advancements and specific applications of DL and CV in detecting defects, ensuring reliability, and maintaining high standards for electronics production.

1) CASE STUDIES: INSPECTION OF PRINTED CIRCUIT BOARDS (PCBs)

One of the critical applications of DL and CV in electronics manufacturing is the inspection of printed circuit boards (PCBs). These technologies enable the detection of many defects, such as missing components, soldering errors, and incorrect component placement, with unprecedented accuracy and speed. For example, Lian et al. [27] developed an automated visual inspection system specifically for PCBs, employing an enhanced Mask R-CNN framework designed for smart city applications. This system demonstrates the adaptability of DL and CV in addressing complex inspection tasks in urban infrastructure. Furthermore, Kim et al. [28] introduced a novel PCB defect detection system that utilized a skip-connected convolutional autoencoder. This approach shows the effectiveness of DL in extracting and learning from intricate patterns in PCB imagery. In addition, Pham et al. [29] proposed a semi-supervised learning model, PCB_SS and a fully supervised model, PCB_FS, which leveraged both labeled and unlabeled images for defect detection. This method highlights the potential of using extensive datasets, including unlabeled images, to train DL algorithms for more precise inspection outcomes. Through these advancements, DL algorithms have become adept at identifying even the subtlest deviations from standard PCB designs. This ensures that every board adheres to stringent quality standards before assembly, significantly enhancing the reliability and performance of electronic products.

2) CASE STUDIES: SEMICONDUCTOR WAFER DEFECT DETECTION

Semiconductor wafers are fundamental to electronic devices, and defect-free production is crucial. DL and CV technologies have been instrumental in identifying defects at the micron level, including scratches, contamination, and pattern irregularities. Wang et al. [30] introduced the Knowledge Augmented Broad Learning System (KABLS), a novel approach for identifying mixed-type defects in semiconductor wafer manufacturing. This system exemplifies the integration of domain knowledge into DL frameworks to enhance defect detection capabilities. To address the challenges specific to semiconductor packaging processes, Kim et al. [31] developed a framework that employs a modified Xception deep learning model to scrutinize defects in the wafer buffer zone. This methodology demonstrates the adaptability of DL models to various stages of semiconductor production. To further refine defect detection strategies, de la Rosa et al. [32] proposed a two-stage approach that merges CV techniques with a lightweight SqueezeNet CNN. This combination provides a balance between the detection accuracy and computational efficiency, which is crucial for inline inspection systems. Complementing these approaches, Limam et al. [33] explored the use of various CNN architectures, including VGG, ResNet, and DenseNet, to automate inline defect detection. By analyzing high-resolution images of wafers during fabrication, DL models offer unparalleled precision in identifying defects, significantly mitigating the risk of semiconductor failures in the final electronic products, thereby upholding the integrity of countless devices that rely on these fundamental components.

3) CASE STUDIES: ENHANCING SURFACE MOUNT TECHNOLOGY PROCESS

Surface Mount Technology (SMT) processes benefit significantly from implementing DL and CV systems. These technologies automate the inspection of SMT lines, ensuring the correct component placement, orientation, and solder paste application. For instance, Zhang et al. [34] introduced the ResNet-34-ECA model, specifically designed to classify welding image defects in SMT. By incorporating data augmentation and the Efficient Channel Attention (ECA) mechanism, this model achieves enhanced defect detection accuracy. Building on the theme of optimizing SMT inspections, Wu et al. [35] developed PCBNet, a tailored lightweight Convolutional Neural Network that focuses on defect inspection within SMT processes. Additionally, Dlamini et al. [36] proposed a novel automatic defect detection system for SMT, utilizing the MobileNetV2 architecture combined with a Feature Pyramid Network (FPN). This system is optimized for real-time identification of mounted devices on Printed Circuit Boards (PCB), showcasing the adaptability of DL models to the needs of high-speed production lines. The ability of DL models to learn from vast amounts of data and improve over time has reduced the SMT-related defects. This evolution not only optimizes the assembly process but also enhances the production yield, underscoring the pivotal role of DL and CV technologies in advancing manufacturing efficiencies.

C. TEXTILE MANUFACTURING

The textile manufacturing sector is witnessing a transformative era by integrating deep learning and computer vision technologies to elevate fabric quality control and inspection processes. This section discusses the application of these technologies for defect detection in textiles and demonstrates their potential to revolutionize the industry.

1) CASE STUDIES: AUTOMATED FABRIC INSPECTION SYSTEMS

Automated fabric inspection systems equipped with DL and CV technologies have significantly improved the detection of fabric defects, such as stains, tears, misweaves, and needle holes. These systems analyze images of fabrics in real time, identifying nearly invisible defects to the human eye. For instance, Jing et al. [37] introduced Mobile-Unet, an efficient convolutional neural network model specifically tailored for fabric defect detection. This model demonstrates the adaptability of CNNs to the complex patterns and textures inherent in fabrics. Further advancing fabric inspection, Hu et al. [38] presented an unsupervised method using an enhanced DCGAN. By incorporating an encoder into the traditional architecture, their approach offers a novel perspective on detecting fabric defects without needing labeled data. Additionally, Cheng et al. [39] proposed SCUNet, a lightweight semantic segmentation network designed to improve defect detection precision, particularly for small defects that pose significant challenges in traditional inspection systems. Convolutional Neural Networks, a cornerstone of DL technologies, have proven highly effective in recognizing complex fabric patterns and distinguishing genuine defects from normal variations in texture. This capability ensures that automated inspection systems can maintain high fabric quality standards, minimizing defects and enhancing the overall production yield.

2) CASE STUDIES: PATTERN RECOGNITION AND CLASSIFICATION

DL models excel in pattern recognition, enabling them to classify fabric defects based on their characteristics, such as shape, size, and severity. This classification is crucial for decision-making processes in textile manufacturing, as it helps determine whether a piece of fabric can be corrected, discarded, or used for lower-quality products. Zhao et al. [40] took this a step further by proposing a VLSTM based integrated CNN model, specifically designed to enhance fabric defect classification. This approach illustrates the potential of combining temporal sequence learning with convolutional neural networks to improve classification accuracy. In parallel, Iqbal Hussain et al. [41] leveraged the powerful ResNet-50 architecture to develop a deep learning model tailored for classifying and recognizing woven fabrics. Their use of data augmentation and transfer learning showcases the adaptability of DL models to the nuanced requirements of fabric pattern analysis. Complementing DL's capabilities, CV technologies play a crucial role by capturing high-resolution images of fabrics. These images serve as the foundational data for DL models, facilitating the precise identification and classification of defects, thereby streamlining the quality assurance process in textile manufacturing.

3) CASE STUDIES: COLOR CONSISTENCY AND DYEING QUALITY CONTROL

Ensuring colour consistency and quality in the dyeing process is another area where DL and CV technologies significantly impact. These technologies can detect uneven dyeing, colour bleeding, and other dye-related defects by analysing fabric images' colour distribution and saturation. For example, Zhang et al. [42] introduced a novel approach to this challenge by proposing an unsupervised learning method utilizing a U-shaped de-noising convolutional auto-encoder (UDCAE). This method, specifically designed for yarn-dyed fabrics, showcases the potential of DL in enhancing the accuracy of defect detection through advanced image analysis techniques. Beyond defect detection, advanced DL algorithms can predict how different fabric types will respond to dyes. This predictive insight allows for preemptive adjustments in the dyeing process, aiming to achieve the desired color quality and consistency across various textiles. Integrating DL and CV into the dyeing quality control process not only enhances the detection of color irregularities but also paves the way for a more informed and adaptable dyeing methodology. This ensures that each piece of fabric not only meets the rigorous standards of color quality but also contributes to the overall aesthetic and durability of the final product.

D. AGRICULTURE MANUFACTURING

Integrating Deep Learning and Computer Vision technologies into agricultural manufacturing processes marks a transformative shift towards precision agriculture and automated farm management. These advancements offer promising solutions to the agricultural sector's most pressing challenges, including crop monitoring, disease detection, and yield prediction, enhancing productivity, sustainability, and resource efficiency.

1) CASE STUDIES: CROP MONITORING AND HEALTH ASSESSMENT

One of the primary applications of DL and CV in agriculture manufacturing is crop monitoring and health assessment. High-resolution images captured by drones or satellites are processed using convolutional neural networks to monitor crop health and growth stages and detect signs of stress or disease. Xiao et al. [43] introduced UAV-Net, a DL-based spatiotemporal fusion (STF) model that enhances crop monitoring precision by generating high-resolution images through the fusion of UAV and satellite data. This model exemplifies the potential of integrating multiple data sources for comprehensive crop surveillance. Expanding on the theme of early detection, Khan et al. [44] developed a robust model employing a Long Short-Term Memory (LSTM) network. This model is tailored for the early identification of staple crops like rice, wheat, and sugarcane in smallholder farms, utilizing time-series Sentinel-2 satellite imagery to facilitate timely agricultural decisions. Moreover, Blekanov et al. [45] presented an innovative approach for assessing the nitrogen status of grain crops using UAV multispectral imagery combined with DL techniques. Their focus on U-Net-based neural network architectures underscores the adaptability of DL models to various aspects of crop health monitoring. The capability of DL and CV systems to provide real-time insights into crop conditions enables the early detection of potential issues, allowing for prompt intervention. This proactive approach significantly reduces the risk of substantial yield loss, underscoring the critical role of these technologies in modern agricultural practices.

2) CASE STUDIES: PRECISION ARGICULTURE

Precision agriculture leverages DL and CV to optimize farming practices based on the analysis of field data. This approach includes the precise application of water, pesticides, and fertilizers tailored to the specific needs of individual plants or field zones. DL algorithms can map field variability by analyzing CV-equipped sensors and aerial imagery data, enabling farmers to implement site-specific crop management practices. Such precision enhances crop yields and minimizes environmental impact by reducing the overuse of agricultural inputs. In this context, Cama-Pinto et al. [46] developed a DL model tailored to improve the performance of wireless sensor networks in greenhouses, addressing the challenge of vegetation interference. This study highlights the critical role of reliable data acquisition in implementing precision agriculture practices effectively. Further advancing precision agriculture, Kumar et al. [47] introduced a multiparameter optimization system that employs Deep Convolutional Neural Networks (DCNN) to refine irrigation planning and scheduling. By providing accurate soil moisture estimations, this system exemplifies the potential of DL in enhancing resource-use efficiency. Additionally, Banerjee et al. [48] explored the application of hybrid deep learning models in classifying banana leaf diseases. Combining CNN with SVM offers a novel approach to plant health management, further contributing to the precision agriculture paradigm. By integrating DL and CV technologies, precision agriculture is setting new standards in farming practices. By enabling site-specific crop management based on detailed field data analysis, these technologies pave the way for sustainable agricultural development, optimizing input use and maximizing productivity.

3) CASE STUDIES: AUTOMATED WEED CONTROL

Weed control is a labor-intensive task that significantly impacts agricultural productivity and costs. DL and CV technologies offer a solution through the development of automated weed detection and removal systems. These systems utilize image recognition algorithms to distinguish between crops and weeds, enabling the precise targeting of weeds with herbicides or mechanical removal methods. This selective approach reduces herbicide usage and preserves the surrounding crop, leading to healthier fields and reduced costs. For example, Moazzam et al. [49] introduced an automated two-stage semantic segmentation approach specifically designed to improve the precision of identifying tobacco plants and weeds in aerial images. This advancement demonstrates the potential of DL in enhancing the accuracy of aerial weed detection. Building on this, Ong et al. [50] explored the use of Convolutional Neural Network classifiers in conjunction with UAV imagery to detect weeds amidst Chinese cabbage crops. Their research highlights the integration of aerial imaging with DL classifiers to refine weed detection in specific crop fields. Furthermore, Tummapudi et al. [51] proposed a novel system that not only detects weeds using deep learning-based object detection algorithms but also incorporates a robotic arm for their physical removal. This system exemplifies the combination of DL with robotics to offer a comprehensive solution for weed management in agriculture. By enabling the precise targeting of weeds, DL and CV technologies reduce herbicide usage and preserve crop health, leading to more sustainable agricultural practices. The development of such automated systems marks a significant stride towards achieving healthier fields and reducing operational costs in agriculture.

4) CASE STUDIES: YIELD PREDICTION AND QUALITY ASSESSMENT

Yield prediction and quality assessment are critical for efficient agricultural manufacturing. DL models analyze crop visual data to estimate yield and assess product quality. This information is invaluable for planning and optimizing the supply chain, from harvest to market. Moreover, quality assessment models can classify agricultural products based on size, colour, and other quality markers, ensuring that only the highest quality produce reaches the market. Such technologies improve market competitiveness and reduce waste by identifying suboptimal products early in the supply chain. Hyper3DNetReg, introduced by Morales et al. [52], is a cutting-edge CNN architecture designed specifically for accurate winter wheat yield predictions, illustrating DL's pivotal role in agricultural resource management and strategic planning. In another study, Tanabe et al. [53] investigated the utility of CNNs for predicting yields through UAV-based multispectral imagery, highlighting DL's transformative potential in enhancing agricultural practices. Additionally, Gururaj et al. [54] created a novel grading system for mangoes that utilizes a synergy of Computer Vision, Deep Learning, and Image Processing techniques. This system aims to automate the mango grading process, making it more efficient and less dependent on manual labor. These innovations emphasize the significant impact of DL and CV technologies in refining yield predictions and quality assessments. By facilitating targeted agricultural interventions, these technologies improve crop yields, superior product quality, and heightened efficiency in agricultural production and distribution processes.

In conclusion, applying DL and CV in agricultural manufacturing is set to revolutionize the sector by enhancing productivity, sustainability, and resource efficiency. As these technologies continue to evolve, their integration into agricultural practices will likely become more widespread, offering innovative solutions to the challenges of modern agriculture.

E. ADDITIONAL SECTORS

Applying Deep Learning and Computer Vision technologies transcends traditional manufacturing domains, offering groundbreaking aerospace, pharmaceuticals, and construction materials solutions. This diversity not only underscores the versatility of DL and CV but also highlights their potential to address quality control challenges across various industries.

1) CASE STUDIES: AEROSPACE MANUFACTURING

In aerospace manufacturing, the stakes for quality and precision are exceptionally high, given the critical nature of components and assemblies. DL and CV technologies are instrumental in ensuring the integrity of parts ranging from turbine blades to fuselage panels. For example, CV systems can detect micro-cracks and corrosion in aircraft components that are difficult for human inspectors to identify. Similarly, DL algorithms are used to analyze X-ray and ultrasound images for internal defects, contributing to the safety and reliability of aerospace components. In this context, Dharmadhikari and Basak [55] delved into the efficacy of two deep neural network (DNN) models specialized in detecting fatigue damage within aerospace-grade aluminum alloys, utilizing ultrasonic testing data. Complementing this approach, Le et al. [56] unveiled a novel nondestructive ultrasonic testing technique powered by a spiking neural network (SNN). This method is particularly designed for pinpointing corrosion in aircraft rivets, showcasing the potential of advanced neural networks in enhancing aerospace component inspections. These advancements underscore the transformative impact of DL and CV in aerospace manufacturing, where they significantly contribute to the meticulous inspection processes required to ensure the safety and reliability of aircraft components.

2) CASE STUDIES: PHARMACEUTICAL INDUSTRY

The pharmaceutical sector benefits immensely from DL and CV in ensuring the safety and efficacy of medications. These technologies are applied in inspecting tablets for correct shape, size, and colour and identifying any physical imperfections that could indicate processing errors. Delving into tablet inspection, Quan et al. [57] unveiled an end-to-end deep learning framework that utilizes ResNet and DenseNet models for the automated detection of defective tablets within pharmaceutical production lines. This innovation demonstrates how DL can enhance quality control in medication manufacturing. Adding to the advancements in this field, Ficzere et al. [58] introduced a cutting-edge Process Analytical Technology (PAT) system. This system melds machine vision with deep learning to offer real-time analysis, classifying tablet coating defects and assessing the thickness of the film coating, thereby ensuring the efficacy and safety of the medication. Moreover, CV systems ensure the integrity of packaging, verifying seal tightness, labeling accuracy, and the presence of safety information. Such meticulous inspection processes are vital in maintaining the trust and health of consumers.

3) CASE STUDIES: CONSTRUCTION MATERIALS

In the construction industry, where a wide array of materials is utilized, Deep Learning and Computer Vision are essential in ensuring the quality and integrity of concrete, steel, and wood. For concrete, CV-based systems are deployed to evaluate the surface texture and structural integrity, identifying potential issues such as cracks or early signs of degradation. To address concrete assessments, Laxman et al. [59] developed an automated system dedicated to the detection and depth estimation of cracks in reinforced concrete structures, leveraging the capabilities of DL for enhanced precision. Similarly, Tabernik et al. [60] crafted SegDecNet++, an innovative DL model for concrete crack detection that uniquely combines pixel-wise segmentation with image-wise classification for comprehensive analysis. Turning to steel, DL models are adept at spotting surface flaws and inconsistencies that might affect the material's structural reliability. Li et al. [61] unveiled a DL model specifically for steel surface defect detection, incorporating a Multiscale Feature Extraction (MSFE) module that utilizes various convolutional kernel sizes for superior feature extraction across different scales. Demir et al. [62] introduced PAR-CNN, a model that enhances defect detection in steel production by integrating parallel training of residual blocks with attention mechanisms for refined classification of surface imperfections. In wood manufacturing, CV technologies are employed to examine grain patterns and identify defects such as knots and splits, critical for ensuring both the material's aesthetic appeal and structural soundness. Lim et al. [63] proposed a compact and efficient CNN model tailored for near real-time wood defect detection, designed to operate on embedded processors. Extending the advancements in wood inspection, Cui et al. [64] introduced CCG-YOLOv7, a model that enhances defect detection on wood floors by integrating novel features such as Center Efficient Layer Aggregation Networks (C-ELAN) and Cascade Center of Gravity Batch Norm (CCG-BN), along with a simplified head module for improved performance. These examples highlight the significant contributions of DL and CV in advancing the construction industry's material quality assurance, showcasing the technologies' pivotal roles in maintaining high standards across various construction materials.

IV. AUTOMATED DEFECT DETECTION TECHNIQUES

The advent of Deep Learning and Computer Vision technologies has significantly advanced automated defect detection across various manufacturing sectors. These technologies enhance the accuracy and efficiency of quality control processes, utilizing sophisticated algorithms to identify defects that may be invisible to the human eye. This section is divided into two subsections, detailing the critical stages in the automated defect detection pipeline: image acquisition and preprocessing techniques, followed by an in-depth exploration of defect detection methods, including supervised, unsupervised, and semi-supervised learning approaches.

A. IMAGE ACQUISITION AND PREPROCESSING TECHNIQUES

The first step in automated defect detection involves capturing high-quality images of the product or material under inspection. This process utilizes specialized cameras and sensors, including digital cameras for visible light imaging, infrared cameras for thermal imaging, and hyperspectral cameras for capturing images across multiple wavelengths. The choice of imaging technology depends on the nature of the defects to be detected and the material properties. Once images are acquired, they undergo preprocessing to enhance their quality and facilitate defect detection. Common preprocessing techniques include:

1) NORMALIZATION

Normalization adjusts the pixel intensity values of an image to a specific range to ensure consistent input for neural networks.

Mathematically, if I represents the original image and I' represents the normalized image, normalization can be defined as:

$$I' = \frac{I - \min(I)}{\max(I) - \min(I)} \tag{1}$$

In this equation, $\min(I)$ and $\max(I)$ represent the minimum and maximum intensity values in the original image, respectively. This transformation scales the intensity values to lie within a predictable range, which is particularly beneficial for neural network models. In the context of surface defect detection in pharmaceutical products, Rački et al. [65] developed a compact deep learning model leveraging convolutional neural networks. A notable aspect of their methodology is incorporating ReLU activation and batch normalization after each convolutional layer, which relies on normalization principles to ensure efficient training and enhance model stability. Similarly, Ouyang et al. [66] proposed a CNN-based automated fabric defect detection system, distinguished by a unique activation layer designed for intricate fabric texture segmentation. Their system applies pre-processing techniques, including normalization via batch normalization after convolutional layers, to improve the model's training stability and operational efficiency.

These examples underscore normalization's vital role in preprocessing images for neural network analysis, demonstrating its effectiveness in diverse applications like pharmaceutical inspection and fabric defect detection.

2) NOISE REDUCTION

Noise reduction aims to smooth out the image to remove random variations in intensity.

a: GAUSSIAN BLUR

Gaussian Blur is a fundamental image pre-processing technique that employs a Gaussian kernel to smooth images, significantly reducing high-frequency noise. This process is crucial for preparing images for more detailed analysis by neural networks. The Gaussian function for a two-dimensional space is given by:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

In this equation, σ represents the standard deviation of the Gaussian distribution, which influences the degree of blurring. A higher σ value results in more pronounced blurring, aiding in the attenuation of noise and minor details in the image. Abayomi-Alli et al. [67] leveraged this technique in their modified MobileNetV2 model for cassava disease recognition. Their innovative approach involved creating synthetic images from high-quality originals by applying various quality reduction techniques, including Gaussian blurring. This method aimed to enhance the model's robustness, particularly for identifying diseases in low-quality images often encountered in real-world scenarios. Similarly, Ashok et al. [68] utilized a Gaussian filter in their CNN model dedicated to detecting and classifying tomato leaf diseases. Applying Gaussian Blur to the leaf images effectively reduced noise and clarified the images, making disease features more distinguishable for the deep learning model. This pre-processing step was instrumental in heightening the model's disease detection accuracy.

These examples underscore the value of Gaussian Blur in image pre-processing for neural networks, showcasing its utility in enhancing data quality and model robustness across various applications.

b: MEDIAN FILTERING

Median Filtering is a non-linear process used in image processing to reduce noise. It replaces each pixel's value with the median value of the intensities found in its immediate neighborhood. Due to its nature, Median Filtering is particularly effective in preserving edges while removing noise, making it a preferred choice for pre-processing in various applications. Unlike linear filters, Median Filtering involves sorting the values in the neighborhood of a pixel and selecting the median value, which does not lend itself to a straightforward mathematical representation. However, its implementation significantly contributes to the clarity and quality of the processed images. In the context of predictive maintenance, Asif et al. [69] demonstrated the utility of Median Filtering in their study. They proposed a model that synergizes Long Short-Term Memory networks with critical preprocessing steps to bolster prediction accuracy for the Remaining Useful Life (RUL) of aircraft turbofan engines. A pivotal aspect of their methodology was employing correlation analysis to discern relevant sensor data, subsequently refined using a moving median filter. This preprocessing step was instrumental in cleansing the data prior to LSTM network training, culminating in enhanced RUL predictions with notably reduced error metrics on the C-MAPSS dataset. Similarly, Alsubai et al. [70] explored the efficacy of Median Filtering in the domain of agricultural disease classification. They introduced a Hybrid DL with an Improved Salp Swarm Optimization-based Multi-class Grape Disease Classification (HDLISSA-MGDC) model. This model incorporates Median Filtering for image preprocessing to eliminate noise from grape leaf images, ensuring the input images are clearer and more conducive to accurate disease classification. The framework utilizes a Dilated Residual Network (DRN) coupled with an Adam optimizer for robust feature extraction, while a CNN-Gated Recurrent Unit (CNN-GRU) ensemble is deployed for classifying grape diseases. The model's performance is further refined using Improved Salp Swarm Optimization (ISSA) to fine-tune the hyperparameters of the CNN-GRU model, achieving superior accuracy and efficiency in grape disease classification.

These instances highlight the pivotal role of Median Filtering in preparing data for complex analyses, showcasing its broad applicability from aerospace engineering to agriculture.

3) ENHANCEMENT

a: HISTOGRAM EQUALIZATION

Histogram Equalization is a fundamental technique in image processing to improve an image's contrast by expanding its intensity distribution. This method is especially valuable in preparing images for neural network analysis by enhancing features that may otherwise be obscured in low-contrast images. The process is mathematically represented as follows:

$$I_{eq} = \left\lfloor \frac{L-1}{MN} \sum_{j=0}^{v} H(j) \right\rfloor \tag{3}$$

In this formula, I_{eq} is the equalized image, M and N denote the image dimensions, v represents the intensity value of a pixel in the original image, and $H(j)$ is the histogram count for intensity level j. This transformation ensures a uniform distribution of pixel intensities, enhancing image contrast. In the realm of automatic optical inspection systems, Chen et al. [71] developed a deep learning model that leverages Histogram Equalization, among other preprocessing techniques, to augment image contrast. This step is crucial for ensuring that the model is trained on images with enhanced features, improving defect detection accuracy in varied lighting conditions and noise levels. Similarly, Khodier et al. [72] implemented an automated inspection system for jacquard-patterned fabrics, employing Contrast-Limited Adaptive Histogram Equalization (CLAHE) to accentuate the fabric's intricate details. This adaptation of Histogram Equalization is particularly effective in addressing the challenges posed by complex fabric patterns, facilitating the CNN model's ability to discern defects with high precision.

These examples underscore the importance of Histogram Equalization in image preprocessing for neural networks, showcasing its utility in enhancing data quality and model robustness across diverse applications.

b: CONTRAST STRETCHING

Contrast Stretching is a vital image pre-processing technique that linearly adjusts the histogram of image pixels, enhancing the contrast by spreading out the most frequent intensity values. The process is mathematically represented as:

$$I' = \frac{(I - \min(I)) \cdot (I_{max} - I_{min})}{\max(I) - \min(I)} + I_{min} \tag{4}$$

In this equation, I' denotes the resulting image after contrast stretching, I is the original image, and I_{min} and I_{max} are the desired minimum and maximum intensity values in the stretched image, respectively. This transformation effectively increases the dynamic range of the image's intensity levels. In the context of apple leaf disease recognition, Rehman et al. [73] developed a parallel real-time processing framework utilizing Mask R-CNN and transfer learning. A key component of their approach is the implementation of a hybrid contrast stretching technique to augment image clarity, a critical factor for precise disease identification. Similarly, Faria et al. [74] introduced a hybrid deep learning framework that combines MobileNet V2 with various recurrent neural network architectures for potato disease classification. They employed contrast stretching to boost the visual quality of images, making the distinction between healthy and diseased tissue more pronounced and thereby facilitating more accurate disease detection.

These instances illustrate the significance of Contrast Stretching in image preprocessing for neural networks, showcasing its utility in enhancing data quality and model effectiveness across diverse applications.

B. DEFECT DETECTION METHODS

In manufacturing, deep learning and computer vision techniques are crucial for enhancing efficiency and accuracy, and they are categorized into supervised, unsupervised, and semi-supervised learning methods. Supervised learning uses labeled data to train models, employing foundational techniques such as Convolutional Neural Networks and specialized versions like ResNet and DenseNet, which address network depth and complexity challenges. Additionally, Recurrent Neural Networks (RNNs), including Long Short-Term Memory networks, are utilized for their ability to process sequences, such as temporal data from assembly lines. On the other hand, techniques such as Autoencoders and Generative Adversarial Networks (GANs) further support feature extraction and image generation tasks. Semi-supervised learning, which combines labeled and unlabeled data, utilizes Deep Belief Networks (DBNs), Semi-Supervised Generative Adversarial Networks (SGAN), and Variational Autoencoders (VAEs) to enhance learning efficiency. These DL and CV methods significantly contribute to optimizing manufacturing processes by improving the automation and precision of tasks. The DL and CV algorithms used in manufacturing are briefly categorized in figure 3.

1) SUPERVISED LEARNING APPROACHES

Supervised learning approaches in the context of automated defect detection involve training models on labeled datasets where the input images are tagged with the correct output, such as the presence or absence of defects. These approaches leverage various algorithms with mathematical underpinnings to learn from the data and predict outcomes for new, unseen images. Below, we delve into some prevalent algorithms and techniques used in supervised learning for defect detection.

Convolutional Neural Networks: Convolutional Neural Networks [75] are a class of deep neural networks most commonly applied to analyzing visual imagery. They have proven highly effective for image recognition, segmentation, and classification tasks. Figure 4 shows the architecture of CNN.

The core operation in a CNN is the convolution operation, which applies a filter (or kernel) to an input image to produce a feature map, highlighting specific features in the image.

Convolution Operation:

At the heart of every CNN is the convolution operation, where a filter, also known as a kernel, is applied to the input image. Given an input image I and a filter K, the convolution operation is defined as:

$$A(x, y) = (K * I)(x, y) = \sum_{m} \sum_{n} K(m, n) \cdot I(x - m, y - n) \tag{5}$$

Here, $A(x, y)$ is the activation map produced by applying the filter K over the input image I. The indices m and n iterate over the kernel, and x and y denote the spatial dimensions of the image.

Feature Learning:

CNNs learn to detect features by adjusting the weights of the filters during the training process. This enables the network to extract meaningful features from the input images, which are then used for classification or other tasks.

Applications of CNNs:

The versatility of CNNs extends beyond basic image processing tasks, as demonstrated in various innovative studies. For instance, Latif et al. [76] explored the application of a modified VGG19 architecture for detecting and classifying rice plant diseases, showcasing the adaptability of CNNs in agricultural contexts. In the semiconductor industry, Chien et al. [77] employed CNNs to automate the identification and categorization of wafer surface defects, highlighting the precision of CNNs in manufacturing quality control. Huang et al. [78] utilized a one-dimensional

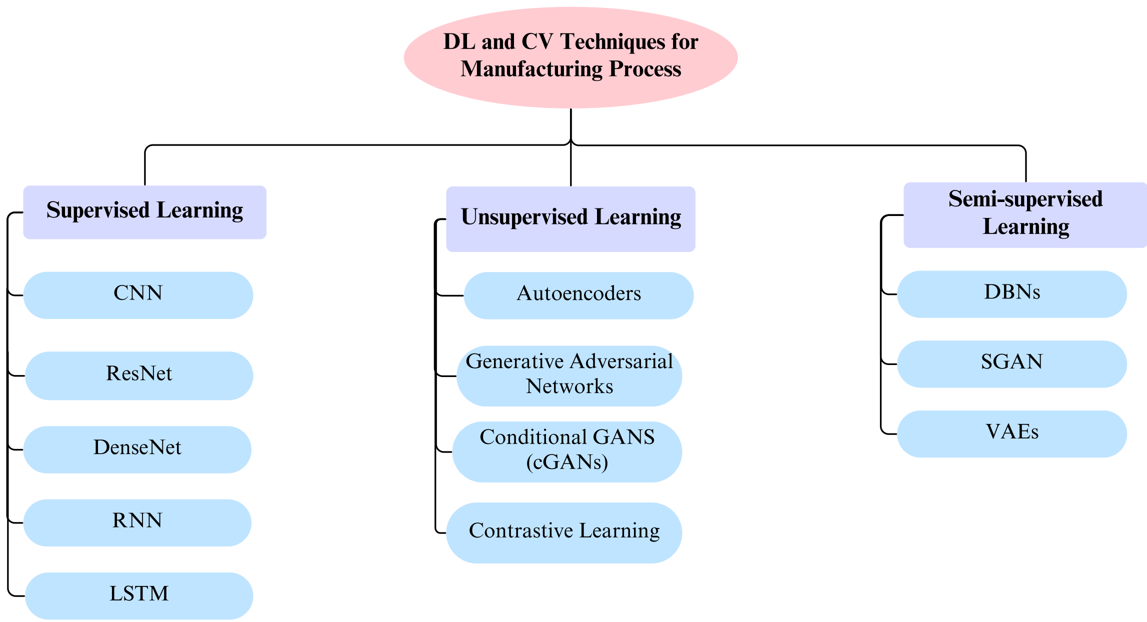

FIGURE 3. The taxonomy of the different DL and CV techniques used for the manufacturing process.

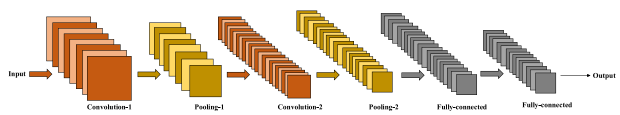

FIGURE 4. Architecture of convolutional neural network.

convolutional neural network (1D-CNN) in conjunction with hyperspectral imaging (HSI) to classify and recognize textile fibers, demonstrating the potential of CNNs in non-destructive testing methods. Focusing on predictive maintenance, Keshun et al. [79] introduced a 3D Attention-enhanced Hybrid Neural Network for estimating the Remaining Useful Life (RUL) of turbofan engines, integrating CNNs to analyze sensor data effectively. In the pharmaceutical domain, Khaouane et al. [80] applied a CNN-based approach to predict plasma protein binding (PPB) of drugs, leveraging CNNs to interpret molecular structures and extract relevant features for further analysis.

These examples underscore the broad applicability of CNNs across various fields, from agriculture and manufacturing to textile analysis and pharmaceutical research. They affirm their status as a cornerstone computer vision and deep learning technology.

Residual Networks:

Residual Networks or ResNets [81], is a type of deep neural network architecture designed to facilitate the training of substantially deeper networks than previously used. The key innovation in ResNets is introducing "skip connections" that allow gradients to flow through the network more effectively during training.

Skip Connections:

The fundamental building block of a ResNet is the residual block, which incorporates skip connections. These connections allow the input of the block to be added to its output, facilitating the learning of an identity function. This is crucial for deep networks to avoid the degradation problem, where accuracy saturates or diminishes with depth. Figure 5 demonstrates the ResNet architecture.

Given an input x, a residual block aims to learn the residual function $F(x)$ concerning the input, rather than a direct mapping. The output of the residual block is:

$$y = F(x) + x \tag{6}$$

Here, $F(x)$ represents the residual mapping to be learned, for multiple layers, F could represent the output of several convolutional layers, with the skip connection adding x directly to the output of these layers.

Identity Mapping by Shortcuts:

The identity shortcut connection, which adds the input x directly to the output of the residual block, requires no additional parameterization and does not introduce extra computational complexity. This simplicity is key to the efficiency and effectiveness of ResNets.

$$y = F(x, \{W_i\}) + x \tag{7}$$

In this equation, $F(x, \{W_i\})$ represents the weighted functions of the layers within the residual block, and x is the input. The addition operation is element-wise.

Applications of ResNets:

ResNets have found wide-ranging applications, demonstrating their adaptability and strength in various fields. In healthcare, Hnoohom et al. [82] refined a ResNet-101 model to boost the precision of blister package identification for hospital medication dispensing, showcasing ResNets'

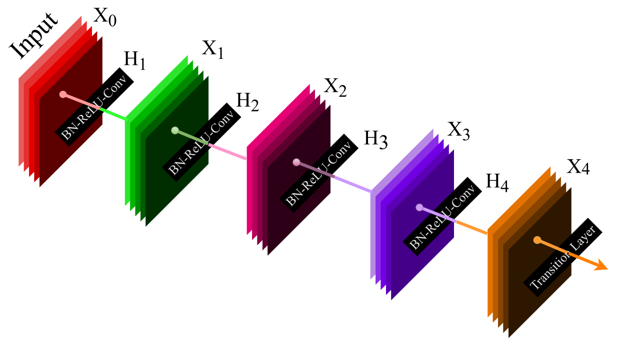

FIGURE 5. The ResNet architecture.

potential in enhancing patient safety through accurate medication management. Tchatchoua et al. [83] applied a 1D ResNet for fault detection in semiconductor manufacturing, analyzing sensor data to pinpoint defects, thus ensuring the quality and reliability of semiconductor products. In agriculture, Stephen et al. [84] utilized pre-trained ResNet models, including enhanced versions with self-attention mechanisms, for diagnosing rice leaf diseases, demonstrating ResNets' contribution to sustainable farming practices. Zhang et al. [85] introduced a Cost-Sensitive ResNet model for PCB defect detection, optimizing the network to classify defects in an imbalanced dataset better, highlighting ResNets' adaptability to industrial quality control. In textile manufacturing, Liu et al. [86] combined Inception and ResNet architectures for fabric image classification, leveraging ResNets' robust feature extraction capabilities to ensure the quality of textile products.

These examples underscore the profound impact of ResNets across diverse sectors, from healthcare and manufacturing to agriculture, by enabling the training of networks that are deeper and more capable than ever before.

Dense Convolutional Networks: Dense Convolutional Networks, or DenseNets [87], are a class of DNNs with a unique architecture where each layer receives input from all preceding layers and passes its feature maps to all subsequent layers. This connectivity pattern promotes feature reuse, significantly reduces the number of parameters, and enhances training efficiency. Figure 6 illustrates the architecture of DenseNet.

Layer Connectivity:

In a DenseNet architecture, the output of the l^{th} layer, denoted as x_l, is derived by aggregating the feature maps of all preceding layers, is computed as follows:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{8}$$

Here, $H_l(\cdot)$ represents a composite function of operations (BN-ReLU-Conv) applied to the concatenated output of all previous layers $[x_0, x_1, \ldots, x_{l-1}]$. Here, x_0 is the input to the first layer.

Feature Concatenation:

The key to DenseNets is the concatenation of feature maps produced by layers before it, which allows the network to propagate features through the network:

$$\begin{aligned} x_l &= H_l([x_0, x_1, \ldots, x_{l-1}]) \\ &= BN - ReLU - Conv([x_0, x_1, \ldots, x_{l-1}]) \end{aligned} \tag{9}$$

This equation illustrates the process of concatenating the outputs of all preceding layers ($[x_0, x_1, \ldots, x_{l-1}]$) and applying the composite function H_l to produce the output of the current layer x_l.

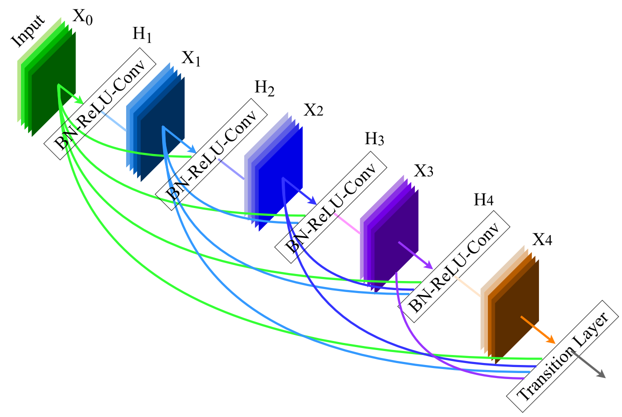

FIGURE 6. The architecture of DenseNet.

Applications of DenseNets:

DenseNets have been successfully applied in various contexts, showcasing their adaptability and efficacy. In the pharmaceutical manufacturing domain, Quan et al. [57] leveraged DenseNet models in an end-to-end framework designed for automatically detecting defective tablets, highlighting DenseNet's applicability in quality control. Similarly, in the agricultural domain, Kulkarni et al. [88] employed DenseNet-264, coupled with the Simple Linear Iterative Clustering (SLIC) segmentation algorithm, to diagnose Coffee Leaf Diseases (CLD). This innovative combination highlights DenseNet's capability to handle complex image data, providing precise and reliable disease detection critical for crop management and protection. Wang et al. [89] introduced an ensemble method that incorporates DenseNet for machinery fault diagnosis, particularly under class-imbalance conditions. This approach demonstrates DenseNet's utility in industrial maintenance. Furthermore, Zhu et al. [90] proposed a fabric defect detection system that adapts DenseNet for edge computing, reducing latency in real-time applications. This adaptation showcases DenseNet's flexibility in addressing computational constraints.

DenseNets offer several advantages over traditional convolutional networks, including reduced parameter count due to feature reuse, improved gradient flow throughout the network, and enhanced feature propagation, making them highly efficient for various tasks in computer vision and beyond.

Recurrent Neural Networks: Recurrent Neural Networks (RNNs) [91] are a class of neural networks designed for processing sequential data. Unlike feedforward neural networks, RNNs have the unique feature of maintaining a hidden state that captures information about the sequence processed so far, making them ideal for tasks like language modeling, time series analysis, and more. The structure of RNN is illustrated in figure 7.

Basic Structure:

At each time step t, an RNN takes an input vector x_t and updates its hidden state h_t based on the previous hidden state h_{t-1} and the current input. The hidden state serves as the network's memory. The RNN then optionally produces an output y_t based on the hidden state. The following equations can summarize the basic operations of an RNN:

$$h_t = \sigma(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \tag{10}$$
$$y_t = W_{hy}h_t + b_y \tag{11}$$

where:
- h_t is the hidden state at time t,
- x_t is the input at time t,
- y_t is the output at time t,
- W_{hh}, W_{xh}, and W_{hy} are the weight matrices,
- b_h and b_y are the bias vectors,
- σ is the activation function, often a non-linear function like tanh or ReLU.

Activation Function:

The choice of activation function σ is crucial for controlling the non-linearity in the model. Common choices include the hyperbolic tangent function (tanh) or the sigmoid function, defined as:

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad \text{(Sigmoid)} \tag{12}$$
$$\sigma(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad \text{(Tanh)} \tag{13}$$

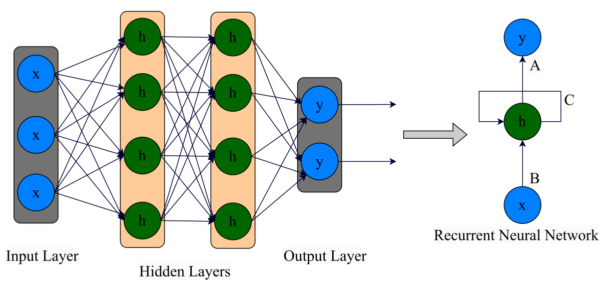

FIGURE 7. The architecture of RNNs.

Applications of RNNs:

The application of RNNs extends across various domains, demonstrating their versatility. Hu et al. [92] utilized an RNN model to predict the quality of cotton yarn, capitalizing on the sequential nature of the yarn production process, a domain where RNN's memory capability offers a significant advantage over models like Multiple Linear Regression and Support Vector Regression. Similarly, Kumari et al. [93] employed RNNs for forecasting banana prices in Gujarat, India, illustrating RNNs' efficacy in time series analysis compared to ARIMA and ANN models. In the pharmaceutical industry, Ruiz Puentes et al. [94] developed PharmaNet, a deep learning architecture incorporating RNNs for predicting active molecules, showcasing the ability of RNNs to process complex sequential data like SMILES strings for drug discovery. Wang et al. [95] applied RNNs for anomaly detection in manufacturing systems' time series data, highlighting RNNs' utility in identifying irregular patterns.

RNNs are powerful models for handling sequential data due to their ability to maintain and update a hidden state across time steps. However, they can suffer from issues like vanishing and exploding gradients, which have led to the development of more advanced variants like LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) networks.

Long Short-Term Memory (LSTM):

Long Short-Term Memory [96] networks are a special subset of RNN, capable of learning long-term dependencies. LSTMs were introduced to overcome the limitations of traditional RNNs, primarily the vanishing and exploding gradient problems, by incorporating a memory cell and three distinct gates: input, output, and forget gates. Figure 8 shows the architecture of LSTM.

Gate Mechanisms:

An LSTM unit updates its cell state and hidden state through a series of gates, each with a specific function:
- **Forget Gate** (f_t): Decides what information to discard from the cell state.
- **Input Gate** (i_t): Decides which values to update.
- **Output Gate** (o_t): Determines the output based on the cell state and the input.

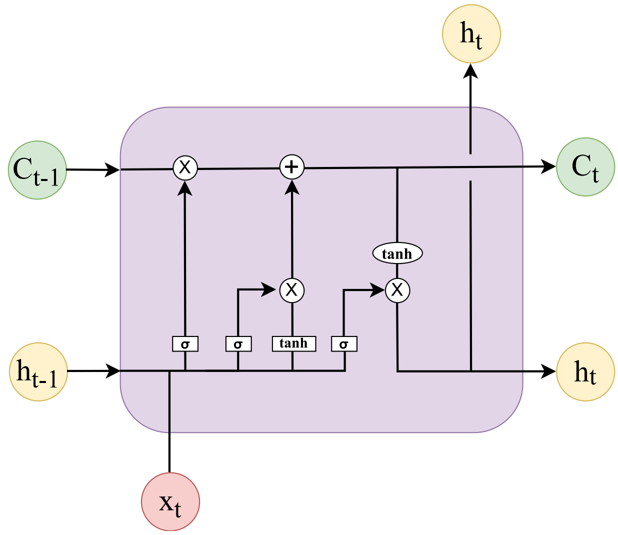

FIGURE 8. The architecture of LSTM.

Applications of LSTMs:

The adaptability of Long Short-Term Memory networks to various domains has been demonstrated through their application to problems requiring an understanding of long-term dependencies. Kumar and Bai [97] utilized LSTMs to address the challenges in textile manufacturing, specifically for fabric texture classification and defect detection. Their work highlights the LSTM's ability to navigate the complexities of sequential patterns in fabric textures, offering a significant improvement in accuracy and efficiency compared to traditional manual inspections. In the field of energy systems, the model introduced by Chung et al. [98], known as the Parallel CNN-LSTM Residual Blocks Attention (PCLRA) model, focuses on anomaly detection within Combined Heat and Power (CHP) engines. This model exemplifies the synergy between CNNs and LSTMs, augmented by residual blocks and attention mechanisms to enhance spatiotemporal feature extraction. This approach underlines the LSTM's effectiveness in isolating and emphasizing critical temporal features from sensor data, vital for maintaining operational reliability in energy systems. The agricultural sector has also benefited from LSTM applications, as seen in the innovative ARIMA-LSTM hybrid model proposed by Ray et al. [99] for forecasting agricultural price series. By integrating ARIMA's and Random Forest's methodologies for input lag selection, this model capitalizes on LSTM's capability to unravel complex temporal dependencies, thereby offering a nuanced approach to predicting price fluctuations in volatile agricultural markets. In healthcare, the attentive LSTM-based framework developed by Qian et al. [100] for predicting Adverse Drug Reactions (ADRs) showcases LSTM's flexibility in sequence-to-sequence modeling. Employing an encoder-decoder structure with attention mechanisms, the model adeptly zeroes in on pertinent input sequence segments, enhancing ADR predictions' precision. This application demonstrates LSTM's potential in healthcare analytics and emphasizes its role in improving patient safety and drug efficacy.

These works illustrate the broad utility of LSTMs across different sectors, from enhancing manufacturing processes and energy system reliability to advancing agricultural forecasting and healthcare analytics. The diverse applications underscore the LSTM's inherent strength in processing and interpreting sequential data, making it a cornerstone technique for tackling complex problems involving temporal dependencies.

Transfer Learning With Pre-Trained Models:

Transfer learning [101] is a machine learning strategy where a model originally created for one task is repurposed as the foundation for a model on a different task. This method is particularly prevalent in deep learning, leveraging pre-trained models to attain cutting-edge performance in scenarios where the dataset is insufficient to train an entirely new model from scratch.

Transfer learning typically involves two main phases:

1) **Pre-training:** A model is trained on a base dataset and task, where it learns general features potentially useful for a range of functions.
2) **Fine-tuning:** The pre-trained model is then adapted to a target dataset and task. This often involves modifying the final layers of the model to suit the target task better and then continuing training on the target dataset.

Mathematical Formulation:

Given a source task T_S with dataset D_S and a target task T_T with dataset D_T, transfer learning aims to improve the learning of the target predictive function $f_T(\cdot)$ using the knowledge from D_S and T_S. The process can be represented as:

$$f_T = \text{fine-tune}(f_S, D_T) \tag{14}$$

Here, f_S represents the model trained on the source dataset D_S.

Benefits of Transfer Learning:

Transfer learning offers several benefits, including:

- **Improved performance:** Pre-trained models can significantly boost the target task's accuracy.
- **Faster convergence:** Since the model is already trained on a related task, converging on the target task requires less time.
- **Reduced data requirement:** Transfer learning can be particularly beneficial when the target dataset is small.

Applications of Transfer Learning:

The utility of transfer learning, especially when combined with pre-trained deep learning models, spans various domains, offering innovative solutions to longstanding problems. Hellapandi et al. [102] explored the potential of transfer learning for the early detection of plant diseases through image classification of diseased plant leaves. Their research involved a comparative analysis of eight pre-trained models, including VGG16, VGG19, ResNet50, InceptionV3, InceptionResnetV2, MobileNet, MobileNetV2, and DenseNet, against a custom-built convolutional neural network. This study underscores the efficacy of leveraging existing deep learning architectures through transfer learning in identifying plant diseases, thereby contributing to early diagnosis and mitigation of crop losses. In the textile sector, Zhu et al. [103] utilized transfer learning with an enhanced ShuffleNetV2 model, CWCNet, to classify cashmere and wool fibers precisely. By incorporating depthwise separable dilated convolution and a novel activation function, EMish's approach significantly improved the accuracy of fiber classification, demonstrating the advantages of transfer learning in refining model performance for specific industry applications. Katsigiannis et al. [104] applied transfer learning to pre-trained deep convolutional neural networks for crack detection on masonry facades. This method, optimized for identifying structural cracks, validates the effectiveness of transfer learning in building inspection tasks, particularly when annotated data is scarce, offering a reliable solution for maintenance and safety assessments. Furthermore, In the pharmaceutical domain, Iwata et al. [105] proposed a classification method for Scanning Electron Microscope (SEM) images of excipients using transfer learning applied to CNNs, such as VGG16 and ResNet50. Their approach achieved high accuracy in discerning excipient types based on SEM images, showcasing the capability of transfer learning to enhance the automatic detection of particle characteristics, which is crucial for the quality evaluation of pharmaceutical raw materials.

These highlight the transformative impact of transfer learning across different fields, from enhancing agricultural disease detection and textile fiber classification to improving structural integrity assessments and pharmaceutical quality control. Researchers and practitioners can improve accuracy and efficiency by adopting pre-trained models and fine-tuning them for specific tasks, even with limited domain-specific data.

Unsupervised Learning Approaches Unsupervised learning does not require labeled data. Instead, it identifies patterns and anomalies in the data based on the assumption that defects are rare and differ significantly from the norm. Autoencoders, for example, are used to learn a compressed representation of normal images. During inference, images with defects will have higher reconstruction errors, indicating the presence of anomalies. This approach is beneficial when collecting labeled data is impractical or when the types of defects are unknown or varied.

Generative Adversarial Networks Generative Adversarial Networks (GANs) [106] represent a category of AI algorithms utilized in unsupervised machine learning, where two neural networks engage in a zero-sum game framework. Ian Goodfellow and colleagues introduced this technique in 2014.

Framework:

GANs comprise of two key components: a generative model, denoted as G, which learns to emulate the data distribution, and a discriminative model, denoted as D, which assesses the likelihood that a sample originated from the training data as opposed to being generated by G. The generator (G) creates fresh data instances, while the discriminator (D) evaluates their authenticity.

Objective Function:

The objective of a GAN can be framed as a minimax game between the generator G and the discriminator D. The discriminator attempts to maximize its ability to correctly classify real and fake samples, whereas the generator tries to minimize this ability. The value function $V(G, D)$ is defined as:

$$\min_G \max_D \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))] \tag{15}$$

Here, x represents a real data instance drawn from the true data distribution $p_{\text{data}}(x)$, z denotes a noise sample from the noise distribution $p_z(z)$, $G(z)$ is the output of the generator given the noise z, and $D(x)$ is the discriminator's estimated probability that x is a real data instance.

Training:

Training a GAN involves optimizing D to maximize $V(D, G)$ for fixed G, and optimizing G to minimize $V(D, G)$ for fixed D. This is done through backpropagation and gradient descent/ascent steps.

Applications of GANs:

Generative Adversarial Networks have found innovative applications across various domains, showcasing their versatility and efficacy. For instance, Li et al. [107] proposed a novel approach for PCB defect detection by generating weak feature defect images using a Generative Adversarial Network (GAN) and improving the Faster R-CNN-based defect detection network. Lamba et al. [108] proposed an optimized classification model for plant diseases using Generative Adversarial Networks. This model employs optimized CNN to classify various plant leaf diseases, with the dataset enhanced through GAN augmentation. In the field of infrastructure maintenance, Xu and Liu [109] utilized GANs to expand training datasets for pavement crack detection, demonstrating GANs' capacity to generate realistic supplementary data that improves the effectiveness of convolutional neural networks.

These applications highlight GANs' transformative potential in generating synthetic data and enriching existing datasets, leading to advancements in automated defect detection, disease classification, and maintenance monitoring.

Conditional GANs (cGANs):

Conditional GANs (cGANs) [110] extend the standard GAN framework by conditioning the generation process on additional information, such as class labels or data attributes. This conditioning allows cGANs to generate data that is not only realistic but also relevant to the specified condition. In the context of quality control in manufacturing, cGANs can be used to generate defect-free samples conditioned on specific attributes or to simulate various defect scenarios for training purposes.

Key Components of cGANs:

- **Generator** (G): The generator takes a random noise vector z and a conditional vector y (which represents the additional information) and generates a sample $G(z|y)$.
- **Discriminator** (D): The discriminator evaluates both real samples x and generated samples $G(z|y)$, along with the conditional vector y, and outputs the probability that the sample is real.

Generator Objective:

The objective of the generator is to increase the likelihood that the discriminator perceives the generated samples as real. This can be formulated as:

$$\begin{aligned}\underset{G}{\text{minimize}}\ \underset{D}{\text{maximize}}\ &\mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\log D(x|y)\right] \\ &+ \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1 - D(G(z|y)|y)\right)\right]\end{aligned} \tag{16}$$

In this formulation, x represents a real data instance drawn from the true data distribution $p_{\text{data}}(x)$, z denotes a noise sample from the noise distribution $p_z(z)$, $G(z)$ is the output of the generator given the noise z, and $D(x)$ is the discriminator's estimated probability that x is a real data instance.

Discriminator Objective:

The discriminator's goal is to accurately classify real and generated samples. The objective function for the discriminator can be expressed as:

$$\begin{aligned}\underset{D}{\text{maximize}}\ &\mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\log D(x|y)\right] \\ &+ \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1 - D(G(z|y)|y)\right)\right]\end{aligned} \tag{17}$$

Here, x denotes a real data instance sampled from the true data distribution $p_{\text{data}}(x)$, z represents a noise sample drawn from the distribution $p_z(z)$, $G(z)$ is the output generated by the generator from the noise z, and $D(x)$ is the probability estimated by the discriminator that x is a real instance.

Applications of cGANs:

cGANs have been effectively applied for quality control in manufacturing processes, particularly in the context of selective laser melting additive manufacturing. In this domain, obtaining diverse data to evaluate internal microstructures is challenging due to the high-speed laser passes over micrometer-scale powder grains. The study by Ramlatchan and Li [111] demonstrated that cGANs could synthesize new, high-quality images that closely match experimental data by learning the underlying features of images corresponding to different laser process parameters. This approach allows for the creation of new data for various parameter combinations, supplementing experimental data and enhancing the evaluation of ideal laser conditions, ultimately supporting more comprehensive quality control and building process monitoring.

2) CONTRASTIVE LEARNING

Contrastive Learning [112] is a powerful unsupervised learning technique where the goal is to learn representations by contrasting similar (positive) and dissimilar (negative) pairs of data points. In unsupervised learning, these pairs are typically created using data augmentation techniques without relying on any labeled data. This approach is particularly useful in domains like quality control in manufacturing, where labeled data may be scarce or expensive to obtain.

Key Concepts in Unsupervised Contrastive Learning

- **Anchor**: A reference data point.
- **Positive Sample**: An augmented version of the anchor or a similar data point.
- **Negative Sample**: A data point that is different from the anchor.

Steps in Unsupervised Contrastive Learning

1) **Data Augmentation**: Generate different views (augmentations) of the same data point to create positive pairs.
2) **Similarity Learning**: Train a model to bring representations of positive pairs closer and push representations of negative pairs apart.

Contrastive Learning Objective

The objective is to learn a representation space where similar data points are close together and dissimilar data points are far apart. This is achieved using a contrastive loss function, such as the InfoNCE (Noise-Contrastive Estimation) loss.

InfoNCE Loss Function

The InfoNCE loss function is defined as follows:

$$\mathcal{L}_{\text{contrastive}} = -\log \frac{\exp(\text{sim}(h(x_i), h(x_i^+))/\tau)}{\sum_{j=1}^{N} \exp(\text{sim}(h(x_i), h(x_j))/\tau)} \tag{18}$$

where:

- $\text{sim}(a, b)$ is the similarity function (e.g., cosine similarity) between representations a and b.
- $h(x)$ is the representation of x learned by the model.
- τ is a temperature parameter.
- N is the number of samples in the batch.

Application in Quality Control:

In quality control for manufacturing processes, contrastive learning can be applied to learn features that distinguish between defect-free and defective products. By training on augmented data without labels, the model can learn robust features that capture the essential characteristics of the products, improving defect detection and classification.

Autoencoders:

Autoencoders [113] are a type of artificial neural network employed for unsupervised learning of effective codings. Their purpose is to acquire a representation (encoding) for a dataset, often utilized for tasks like dimensionality reduction or feature learning. The architecture of autoencoders is illustrated in figure 9.

Architecture:

An autoencoder comprises two primary components: the encoder and the decoder. The encoder condenses the input into a latent-space representation, while the decoder reconstructs the input from this latent space. Formally, if we consider an input vector $x \in \mathbb{R}^n$, the encoder and decoder functions can be represented as:

$$h = f(x) = \sigma(Wx + b) \tag{19}$$

$$\hat{x} = g(h) = \sigma'(W'h + b') \tag{20}$$

where h is the encoded representation (hidden layer), $\hat{x}$ is the reconstructed input, σ and σ' are activation functions, and W, W', b, b' are the parameters of the encoder and decoder, respectively.

Objective Function:

Training an autoencoder aims to minimize the difference between the input x and its reconstruction $\hat{x}$. This is often measured using a loss function such as the mean squared error (MSE):

$$L(x, \hat{x}) = ||x - \hat{x}||^2 = \sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \tag{21}$$

Learning Process:

The parameters of the autoencoder (W, W', b, b') are learned by using an optimization algorithm (e.g., stochastic gradient descent) to minimize the loss function. This process entails calculating the gradient of the loss function with respect to the parameters and adjusting the parameters in a manner that diminishes the loss.

Applications of Autoencoders:

Autoencoders have been adeptly applied across various sectors, showcasing their utility in unsupervised learning tasks. For instance, Chow et al. [114] utilized convolutional autoencoders for anomaly detection in concrete structures, a critical aspect of infrastructure monitoring, by training the model to identify deviations from the norm, indicating defects. In the marine industry, Qu et al. [115] combined Echo State Networks with Deep Autoencoders to enhance

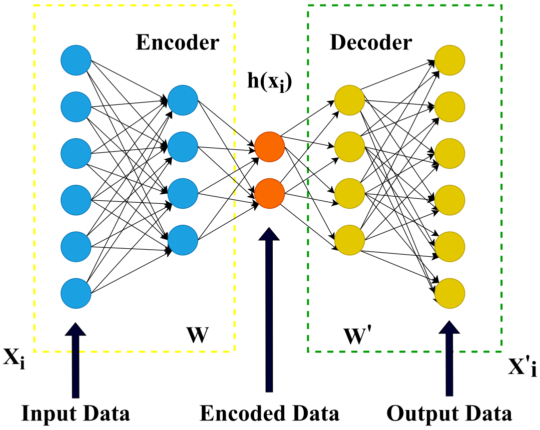

FIGURE 9. The architecture of autoencoders.

predictive maintenance of marine diesel engines by detecting anomalies in sensor data sequences. Furthermore, in the agricultural domain, Y1lmaz et al. [116] integrated Stacked AutoEncoders with CNNs for assessing lemon quality, demonstrating autoencoders' effectiveness in preprocessing and feature extraction for classification tasks.

These diverse applications underscore autoencoders' versatility in extracting meaningful representations from complex datasets, facilitating anomaly detection, predictive maintenance, and quality assessment across different fields.

3) SEMI-SUPERVISED LEARNING APPROACHES

Semi-supervised learning represents a hybrid machine learning approach that is partially supervised and partially unsupervised. It combines a small amount of labeled data with many unlabeled data to train models. This approach is beneficial when acquiring labelled data, which is expensive or time-consuming. Semi-supervised learning algorithms, such as self-training and co-training, leverage the labeled data to learn initial models and then iteratively refine these models by incorporating unlabeled data. This method can improve the model's generalizability and ability to detect novel defects. Figure 10 illustrates the workflow diagram of semi-supervised learning.

Integrating DL and CV in automated defect detection represents a significant leap forward in manufacturing quality control. By employing a combination of image acquisition and preprocessing techniques alongside sophisticated defect detection methods spanning supervised, unsupervised, and semi-supervised learning approaches, manufacturers can achieve higher accuracy, efficiency, and reliability in identifying and addressing defects across various sectors. As these technologies evolve, their application scope will expand, further revolutionizing industrial product management and quality assurance practices.

C. ANALYSIS OF DEFECT DETECTION METHODS

In this subsection, we present a detailed overview of various defect detection methods used in manufacturing processes, categorized into supervised learning, unsupervised learning, and semi-supervised learning. Each method is evaluated based on its advantages, disadvantages, and specific application domains to provide a comprehensive understanding of their suitability for quality control in industrial settings.

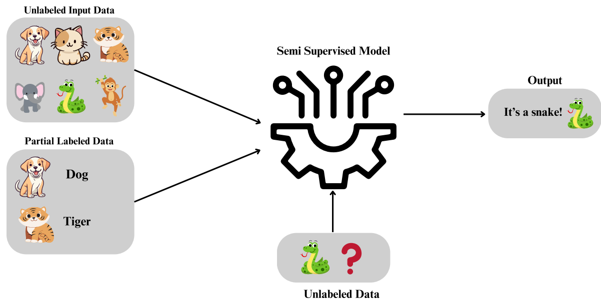

FIGURE 10. Semi-supervised learning workflow.

Table 3 provides a comprehensive overview of various defect detection methods categorized into supervised, unsupervised, and semi-supervised learning algorithms. It highlights the specific algorithms within each category, such as CNNs, ResNet, DenseNet, RNNs, and LSTMs for supervised learning; GANs, cGANs, Contrastive Learning, and Autoencoders for unsupervised learning; and DBNs, SGANs, and VAEs for semi-supervised learning. Each algorithm's advantages, such as high accuracy, effective representation learning, and dimensionality reduction, are balanced against their disadvantages, including the need for large amounts of labeled data, computational intensity, and training complexity. The table also outlines specific application domains, emphasizing their use in tasks like defect detection, synthetic data generation, anomaly detection, and predictive maintenance. This structured information helps in selecting the appropriate method for enhancing quality control in manufacturing processes by considering their strengths, limitations, and suitable applications.

V. RECENT SOTA WORKS

A. AUTOMOTIVE MANUFACTURING

In the automotive manufacturing sector, adopting deep learning and computer vision is revolutionizing quality assurance processes, establishing unprecedented standards for accuracy and efficiency. This section discusses the significant impact of these cutting-edge technologies on producing automotive parts and vehicles. These technologies are pivotal in driving industry advancements by enhancing product quality and ensuring the consistency of manufacturing operations. Table 4 illustrates some of the latest deep learning developments and computer vision in automotive manufacturing.

B. ELECTRONICS MANUFACTURING

In the electronics manufacturing industry, deploying deep learning and computer vision is reshaping quality assurance practices, setting unprecedented standards for accuracy and productivity. This section examines the profound effects of these advanced technologies on the fabrication of electronics,

TABLE 3. Summary of algorithms in deep learning and computer vision for quality control.

Algorithm Type	Algorithm Name	Advantages	Disadvantages	Specific Application Domain
Supervised Learning	CNN	High accuracy in image recognition, spatial invariance	Requires large amounts of labeled data, computationally intensive	Defect detection, image classification
	ResNet	Solves vanishing gradient problem, deeper networks possible	More complex and computationally expensive	Detailed inspection tasks, image recognition
	DenseNet	Alleviates vanishing gradient, improves feature propagation	High memory consumption	Quality control, complex defect detection
	RNN	Good for sequential data, captures temporal dependencies	Difficulty in training long sequences due to vanishing gradients	Sequence prediction, time-series analysis
	LSTM	Addresses vanishing gradient in RNNs, effective for long-term dependencies	Computationally intensive, longer training times	Predictive maintenance, defect prediction
Unsupervised Learning	GANs	Generates high-quality synthetic data, unsupervised training	Training instability, mode collapse	Defect simulation, synthetic data generation
	cGANs	Controlled data generation with conditional inputs	Complex training process, requires careful design	Customized defect generation, specific quality control tasks
	Contrastive Learning	Effective representation learning without labels	Requires large batch sizes, complex training	Anomaly detection, feature extraction
	Autoencoders	Dimensionality reduction, noise reduction	Risk of overfitting, may lose important information	Anomaly detection, feature compression
Semi-supervised Learning	DBNs	Combines supervised and unsupervised learning, good feature extraction	Difficult to train, computationally expensive	Feature learning, hybrid quality control methods
	SGAN	Leverages limited labeled data with GANs	Complex training, mode collapse	Semi-supervised defect detection, limited data scenarios
	VAEs	Generates diverse data, probabilistic approach	Blurry outputs, complex training	Defect simulation, probabilistic quality control

including the assembly of components and entire devices. These tools play a crucial role in elevating product quality and solidifying the robustness of production lines. Table 5 highlights some key recent developments in deep learning and computer vision in electronics manufacturing.

C. TEXTILE MANUFACTURING

In textile manufacturing, integrating deep learning and computer vision technologies is transforming quality control operations, establishing new benchmarks for accuracy and efficiency. This section explores the significant impact of these innovations on the production of textile machinery and equipment and the processing of fabrics and garments. It highlights how these technologies improve product quality and ensure the dependability of manufacturing workflows. Table 6 summarizes recent advancements in deep learning and computer vision within the textile manufacturing sector.

D. AGRICULTURE MANUFACTURING

In the realm of agricultural manufacturing, the application of deep learning and computer vision techniques is revolutionizing quality control processes, setting new standards for precision and efficiency. This section delves into the transformative impact of these technologies on the manufacturing of agricultural machinery, equipment, and the processing of agricultural produce, highlighting their role in enhancing product quality and ensuring the reliability of the manufacturing processes. Table 7 summarizes recent contributions to DL and CV in agricultural manufacturing.

E. OTHERS SECTORS

DL and CV are also driving innovations in a range of industrial sectors beyond agriculture, significantly elevating quality control measures. In aerospace manufacturing, these technologies are instrumental in the meticulous assembly of aircraft components, ensuring unparalleled precision and safety standards. The pharmaceutical industry benefits similarly, with DL and CV facilitating the rigorous inspection of drug manufacturing processes to guarantee product purity and compliance with health regulations. Additionally, in the construction materials sector, these advanced techniques are applied to monitor the quality of materials such as cement and steel, enhancing structural integrity and reliability. Each of these sectors showcases the versatile and transformative potential of DL and CV in modern manufacturing environments. Table 8 summarizes some of the recent contributions of DL and CV in these additional sectors.

VI. EMERGING TRENDS AND FUTURE DIRECTIONS

The landscape of industrial product management, particularly in the realms of quality control and defect detection, is undergoing a transformative shift driven by rapid technological advancements and the integration of artificial intelligence. This section explores the future directions and emerging trends set to redefine these processes' efficacy and efficiency.

A. ADVANCES IN SENSOR TECHNOLOGY AND IMAGING TECHNIQUES

The future of automated defect detection in manufacturing hinges significantly on advancements in sensor technology

TABLE 4. Recent advancements in deep learning and computer vision for automotive manufacturing.

Ref	Dataset	Pre-processing Techniques	Model	Result	Limitation
Wang *et al.* [14]	Custom Dataset	Data Augmentation	ResNet-50 + Transfer Learning	Accuracy: 94.4%	The accuracy of defect identification heavily depends on the performance of the trained deep learning model, specifically the revamped ResNet-50. If the model is not well-trained or lacks generalizability, it might not effectively identify or classify various types of defects, particularly in diverse environments.
Yu *et al.* [15]	Custom Dataset	Labeling, Resizing	AMCD based on YOLOv3-tiny	mAP: 84.96%	The model may be limited by MAVs' on-board computational capabilities, which could restrict its deployment in scenarios requiring intensive computational power.
Upadhyay *et al.* [17]	Custom Dataset	Resizing, Denoising, Labeling, Data Augmentation	Customized U-Net + GANs	Accuracy: binary cross-entropy loss: 99.5% focal loss: 99.5% Dice Coefficient: Jaccard loss: 88.92% Tversky loss: 94.4%	The model is dependent on synthetic data, which might not adequately represent real-world variability, potentially reducing its effectiveness in practical applications.
Wang *et al.* [18]	Custom Dataset	Cropping, Data Augmentation	DBFF-YOLOv4	mAP: 99.58%, Recall: 91.87%	The proposed model might face a potential decrease in performance when detecting very small defects or in different imaging conditions not represented in the training set.
Zubayer *et al.* [117]	Custom Dataset	Normalization, Resizing, RGB Conversion, Data Augmentation	YOLOv8	mAP: 99.5%	The model's training was limited to a relatively small dataset derived from 151 original images, which might not provide sufficient variability to fully assess the model's generalizability and effectiveness across more diverse real-world conditions.
Ma *et al.* [118]	Custom Dataset (ABSDD), KSDD2, DAGM	Semantic Prior Mining (SPM), Defect Enhancement Perception (DEP)	SPDP-Net	ABSDD: Precision: 95.9%, KSDD2: AP: 94.1%, DAGM: Accuracy: 100%	While the model performs strongly in detecting tiny and weak defects on aero-engine blades, this effectiveness may be constrained when applied to environments or defect types that were not adequately represented during training.

and imaging techniques. These technologies are pivotal in capturing minute details and nuances of a product, critical for accurate defect detection and quality control. The evolution in this area enhances image resolution and expands the data types that can be analyzed, encompassing a broad spectrum of electromagnetic wavelengths and imaging modalities. Here are some key developments and their implications:

- High-Resolution Imaging: Modern sensors are achieving higher resolutions, allowing for the detection of subtler defects that were previously indiscernible. This enhancement, in detail, facilitates more accurate defect classification and localization, enabling manufacturers to maintain higher quality standards.
- Multispectral and Hyperspectral Imaging: These imaging techniques capture data across multiple electromagnetic spectrum bands, providing a more comprehensive analysis of materials than visible light imaging alone. They are instrumental in identifying material inconsistencies and anomalies not visible to the naked eye, such as in the food, pharmaceutical, and agricultural sectors.
- 3D Imaging and Laser Scanning: Three-dimensional imaging technologies offer detailed surface and volumetric information about products, including laser scanning and structured light. This capability is crucial for detecting defects that involve shape or volume deviations, such as warping or internal voids, which are challenging to identify with traditional 2D imaging.
- Thermal Imaging: Thermal cameras detect heat signatures and object variations, indicating electrical faults, mechanical wear, or other issues that may not be visible in the light spectrum. This technology is invaluable for predictive maintenance and ensuring the safety and reliability of electronic and mechanical components.
- High-Speed Imaging: Advances in sensor and processing speeds allow for high-speed imaging, enabling the inspection of products on fast-moving production lines without compromising image quality. This capability is essential for maintaining throughput in high-volume manufacturing environments while meeting quality standards.
- Embedded Vision Systems: Integrating compact, robust vision systems directly into manufacturing equipment and robots. This allows for real-time, on-the-fly inspections and adjustments, further automating the manufacturing process and reducing the latency between defect detection and correction.
- Machine Vision with Artificial Intelligence: Combining advanced imaging technologies with AI and deep learning enables the development of systems that can learn from data to improve defect detection over time.

TABLE 5. Recent advancements in deep learning and computer vision for electronics manufacturing.

Ref	Dataset	Pre-processing Techniques	Model	Result	Limitation
Pham *et al.* [29]	Custom Dataset	Resizing, Data Augmentation	PCB_SS & PCB_FS	Accuracy: PCB_SS: 92.27% PCB_FS: 89.67, Recall: PCB_SS: 94% PCB_FS: 90%	The PCB_SS model necessitates region-of-interest (ROI) images from an AFVI system to identify defect candidates, limiting its direct applicability to entire PCBs or assembled PCBs without additional preprocessing to acquire suitable patch images.
Wang *et al.* [30]	MixedWM38	Standardization	KABLS	Accuracy: 99.3%	The effectiveness of the selective sampling mechanism provided by BSSM may be confused by overlapped defects in mixed-type patterns containing three or four single-type defects, such as the difficulty in distinguishing between a donut defect and a center defect in some patterns.
de la Rosa *et al.* [32]	Custom Dataset	Resizing, Cropping, Data Augmentation	Lightweight SqueezeNet CNN	Accuracy: 99.4%, F1-score: 99.4%	The study addressed the data imbalance challenge by applying data augmentation, however, there remains a potential limitation in the representation and variability of these synthetic images, which might not fully capture the complexity and diversity of real-world defects in semiconductor manufacturing processes.
Zhang *et al.* [34]	Custom Dataset	Cropping, Data Augmentation, Standardization and Normalization	ResNet-34-ECA	Accuracy: 98.2%	The proposed model is not tested on datasets from other domains. Further exploration needed with optimizers and loss functions. Relies on datasets provided by Liuzhou United Automotive Electronics, which may limit generalizability.
Dlamini *et al.* [36]	Custom Dataset	Histogram Equalization, Bilateral Filtering, Data Augmentation, Labeling	MobileNetV2 with FPN	Precision: 97.92%, Recall: 96.25%, F1-score: 97.08%	The study's sample collection was limited to a local company, resulting in poor control over data collection and a dataset that may not encompass the full diversity of SMT defects.
Kaya, Gulhan Ustabas [119]	Custom Dataset	Phase Retrieval Process, Binarization	CNN	Accuracy: 99%	The proposed study might be sensitive to noise and imaging conditions, potential for overfitting in the CNN model, and the challenge of generalizing the model to PCBs with different designs or defect types not included in the training dataset.

These systems can adapt to new defect types and variations without explicit programming, enhancing their versatility and effectiveness.

These advancements in sensor technology and imaging techniques are not just elevating the standards of quality control but are also making the inspection processes more adaptable, efficient, and cost-effective. As these technologies continue to evolve and converge with other digital innovations, they will play a pivotal role in shaping the future of manufacturing, ensuring that it is more resilient, sustainable, and aligned with the demands of modern markets.

B. THE ROLE OF BIG DATA AND ANALYTICS IN DEFECT DETECTION

In the era of Industry 4.0, integrating Big Data and analytics into defect detection processes represents a significant leap toward predictive maintenance and real-time quality control. When effectively harnessed, the vast volumes of data generated by manufacturing operations offer a goldmine of insights for improving product quality and operational efficiency.

- Enhanced Predictive Capabilities: Big Data analytics enables the aggregation and analysis of diverse data sources, including historical quality control records, machine performance data, and real-time production metrics. By applying advanced analytics and machine learning algorithms to this data, manufacturers can identify patterns and correlations that were previously undetectable. This predictive insight allows for anticipating potential defects before they occur, shifting the focus from reactive to preventive quality control measures.
- Real-time Monitoring and Response: The real-time production data analysis through Big Data technologies facilitates immediate feedback loops. Sensors and IoT devices continuously monitor the manufacturing process, detecting anomalies and deviations from the norm. By integrating this data with deep learning and computer vision systems, manufacturers can instantly identify and rectify defects, minimizing waste and reducing downtime.
- Customization and Adaptation: Big Data analytics also supports customizing defect detection systems to specific manufacturing environments and product types. Machine learning models can be trained on vast datasets to recognize complex defect patterns unique to particular materials, components, or manufacturing techniques. This level of customization enhances the accuracy and relevance of defect detection efforts.
- Quality Improvement Through Data Integration: Integrating data across the manufacturing ecosystem offers a holistic view of the quality control process. Analytics

TABLE 6. Recent advancements in deep learning and computer vision for textile manufacturing.

Ref	Dataset	Pre-processing Techniques	Model	Result	Limitation
Jing *et al.* [37]	Yarn-dyed Fabric Images & Fabric Images	Data Augmentation, Labeling	Mobile-Unet	IOU: Yarn-dyed: 92% Fabric Images: 70% Recall: Yarn-dyed: 92% Fabric Images: 80%	Class imbalance challenge due to the rarity of fabric defects compared to normal samples, which complicates the training of deep learning models.
Cheng *et al.* [39]	AITEX	Image Cropping, Data Enhancements, Data Augmentation	SCUNet	Accuracy: 98.01%, Recall: 96.81%, Specificity: 98.07%	The study acknowledges limitations such as the use of grayscale images only and the requirement for fabrics to be flat, which may restrict the application to a variety of color textiles and irregular planes. Their method predicts more backgrounds as defects when segmenting the edges of defects.
Zhang *et al.* [120]	FDI-1000, DHU-Semi1000 & Aliyun-Semi10500	Resizing, Transformation	SSA-ULNet	Accuracy: FDI-1000: 87.62% DHU-Semi1000: 86.57% Aliyun-Semi10500: 62.82 %	The study outlined the requirement for significant computational resources due to the deep structure of the model. Additionally, the model may face challenges in situations where there are not enough labeled samples for effective training, which is a common issue in supervised and semi-supervised deep learning approaches.
Alruwais *et al.* [121]	Fabric defect database	Contrast Limited Adaptive Histogram Equalization (CLAHE)	HMFODL-FDD	Accuracy: 95.47%, Sensitivity: 92.97%, Specificity: 96.48%	A notable limitation of the proposed study is the relatively small size of the dataset, which may compromise the model's ability to accurately recognize and categorize defects in fabric types or variations not included in the initial training dataset.
Revathy, G and Kalaivani, R [122]	HKBU database	CLAHE	Improved Mask RCNN	Accuracy: 97.8%, Precision: 94.37%	Although the model shows high accuracy with the HKBU dataset, which includes specific types of patterned fabrics, it might not perform as effectively on fabrics with different patterns or more complex textures that were not included in the training dataset.

can uncover insights from the production line and supply chain operations, customer feedback, and post-market performance. This comprehensive approach enables manufacturers to identify the root causes of defects more effectively and implement systemic improvements.

- Operational Efficiency and Cost Reduction: Manufacturers can achieve significant operational efficiencies by leveraging Big Data and analytics for defect detection. Predictive maintenance reduces manual inspections and repair downtime, while real-time monitoring helps maintain optimal production conditions, minimizing waste. Additionally, the insights gained from data analytics can inform strategic decisions regarding process improvements, resource allocation, and investment in technology upgrades.

In conclusion, the role of Big Data and analytics in defect detection is transformative, offering enhanced capabilities for identifying and preventing defects and opportunities for operational optimization. As these technologies evolve, their integration into manufacturing processes will become increasingly sophisticated, driving further product quality and manufacturing efficiency improvements.

C. THE POTENTIAL OF REINFORCEMENT LEARNING

Reinforcement Learning (RL) presents a promising avenue for revolutionizing industrial product management by enabling autonomous decision-making in dynamic environments. Unlike supervised and unsupervised learning, which rely on labelled datasets or predefined objectives, RL acquires decision-making skills through interacting with the environment and receiving feedback in the form of rewards or penalties.

- Adaptive Decision-Making: RL algorithms can adaptively learn optimal decision policies by continually interacting with the manufacturing environment. This adaptability is crucial when manufacturing conditions constantly change, allowing RL agents to adjust their real-time strategies to maximize efficiency and productivity.
- Autonomous Optimization: RL techniques can optimize various aspects of industrial processes, such as resource allocation, scheduling, and equipment maintenance. RL agents can learn to optimize production parameters and minimize defects by exploring different actions and evaluating their long-term consequences.
- Complex Control Systems: RL enables the development of sophisticated control systems to handle complex manufacturing processes. RL agents can learn to precisely control robotic arms, automated inspection systems, and other machinery, leading to improved product quality and throughput.
- Adaptive Fault Detection: RL algorithms can enhance fault detection systems by continuously learning from operational data and adapting to changes in the production environment. RL agents possess the capability to identify abnormalities and deviations from

TABLE 7. Recent advances in deep learning and computer vision for agricultural manufacturing.

Ref	Dataset	Pre-processing Techniques	Model	Result	Limitation
Mallick *et al.* [123]	Custom Dataset	Resizing, Data Augmentation	MobileNetV2 with Transfer Lerning	Accuracy: 93.65%, Precision: 92.6%, Recall: 94.4%	The model's generalization was limited due to training on a small dataset, expanded through augmentation, necessitating more real-world images for robustness.
Yang *et al.* [124]	Plant Village	Resizing, Patch Embedding	Multi-Path Context Feature Aggregation Network	Accuracy: 99.5%	Limited generalizability due to focus on one dataset. Further testing on diverse datasets is needed for robustness and suitability confirmation in various agricultural contexts.
Dash *et al.* [125]	Plant Village	Cropping, Data Augmentation	DenseNet201 + SVM	Accuracy: 94.6%	The evaluation of the proposed model was limited to only four maize leaf disease classes. Further testing across additional classes is required for comprehensive validation.
Roy *et al.* [126]	Plant Village	Annotation, Data Augmentation	PCA DeepNet	Accuracy: 99.6%, Precision: 98.55%, Recll: 98.49%	Limited to tomato leaf diseases, further reduction in CNN layers suggested for efficiency.
Zhang *et al.* [127]	Custom Dataset	Image Enhancements, Transformation, Data Augmentation	AlexNet	Sorting Accuracy: 89%	The study identifies issues related to the time and efficiency of the sorting process, highlighting these as key areas for future enhancements.
Padmapriya *et al.* [128]	Custom Dataset	Filtering, Data Augmentation	Multi-Stacking Ensemble Model	Accuracy: 98.96%, Precision: 96.14%, Recall: 99.65%	Limited epoch count for proposed model and potential result fluctuations due to diverse soil samples and hyperparameter configurations.
Li *et al.* [129]	Plant Village & Custom Dataset	Resizing, Data Augmentation	MDCDenseNet	Accuracy: 98.84%	Extended model training duration due to a high parameter count, with plans for future optimization efforts.

normal operations, facilitating preemptive maintenance and reducing downtime.

- Exploration of Novel Solutions: RL encourages exploration and experimentation, allowing industrial systems to discover novel solutions and optimize processes beyond human intuition. By iteratively improving decision-making through trial and error, RL agents can uncover efficient strategies that may not be apparent through traditional methods.
- Integration with Digital Twins: RL techniques can be integrated with digital twins, virtual replicas of physical manufacturing systems, to simulate and optimize production processes before implementation in the real world. This integration enables safer experimentation and facilitates rapidly deploying optimized control policies.

In conclusion, the potential of Reinforcement Learning in industrial product management is vast and multifaceted. By enabling adaptive decision-making, autonomous optimization, and the exploration of novel solutions, RL promises to revolutionize the efficiency, quality, and resilience of industrial processes in the era of Industry 4.0. As research in RL continues to advance, its application in industrial settings is expected to become increasingly prevalent, driving innovation and transforming manufacturing paradigms.

D. INTEGRATION OF DL AND CV TECHNOLOGIES WITH OTHER INDUSTRY 4.0 COMPONENTS

The Industry 4.0 paradigm, characterized by integrating digital technologies into manufacturing processes, profoundly synergizes with Deep Learning and Computer Vision technologies. This integration heralds a new era of intelligent manufacturing, enhancing efficiency, productivity, and adaptability. Below are some detailed insights into how DL and CV technologies are merging with other Industry 4.0 components:

- Internet of Things (IoT): DL and CV technologies seamlessly integrate with IoT devices to create intelligent sensor networks throughout manufacturing facilities. These IoT devices capture real-time data, such as temperature, pressure, and vibration, providing valuable inputs to DL and CV models for predictive maintenance, quality control, and process optimization.
- Robotics and Automation: DL-powered robotic systems with CV capabilities are revolutionizing industrial automation. These robots can handle complex tasks such as object recognition, grasping, and manipulation with unparalleled accuracy and efficiency. Furthermore, collaborative robots, or cobots, are emerging as critical players in human-robot collaboration scenarios, enhancing flexibility and safety in manufacturing environments.
- Digital Twins: DL and CV technologies are pivotal in developing and deploying digital twins—virtual replicas of physical assets, systems, and processes. By continuously monitoring and analyzing data from physical assets, digital twins enable predictive maintenance, performance optimization, and scenario simulation. DL algorithms process vast amounts of data from sensors and cameras to represent real-world phenomena accurately.

TABLE 8. Recent advancements in deep learning and computer vision for other sectors.

Ref	Dataset	Pre-processing Techniques	Model	Result	Limitation
Le *et al.* [56]	Custom Dataset	Short-Time Fourier Transform (STFT)	Spiking Neural Network (SNN)	Accuracy: 95.4%	The challenge of detecting corrosion using SNN due to the potential deformation of the water membrane during scans, which can introduce noise and affect the model's accuracy.
Laxman *et al.* [59]	Public Dataset by [130] & Custom Dataset	Cropping	binary-class CNN	Accuracy: Public Dataset: 99.9% Custom Dataset: 93.7%	The models were trained to detect and predict crack depths only under monotonic loading conditions. They were also limited by the number of images and controlled lighting conditions used in the dataset.
Tabernik *et al.* [60]	Public Dataset by [131]	Data Augmentation	SegDecNet++	Dice score: 81%, IoU: 71%	The proposed model faces difficulty in accurately segmenting thin cracks and cracks on dark asphalt with soft characteristics. This limitation suggests that the model may struggle with certain crack types or surface conditions.
Li *et al.* [61]	NEU-DET	Resizing	YOLOv5s	mAP: 73.08%	The method might face challenges in generalization across different types of steel surfaces or under varying operational conditions not represented in the NEU-DET dataset.
Lim *et al.* [63]	Public Dataset by [132]	Resizing, Data Augmentation	YOLOv4-Tiny	mAP: 94.53%	The aggressively pruned model might face limitations in detecting more subtle wood defects due to the reduced complexity and potentially less ability to generalize across different wood types or less controlled environments. There's a trade-off between model size and depth of feature learning that could affect performance in practical applications outside the controlled conditions of the dataset.
Cui *et al.* [64]	Custom Dataset	Rescaling, Data Augmentation	CCG-YOLOv7	mAP: 94.8%, Precision: 92.7%	The model's effectiveness under variable industrial conditions, such as different lighting intensities, diverse wood patterns, and complex defect types, may require further validation to ensure robustness and generalizability across different production settings.
Wang *et al.* [133]	Custom Dataset	Normalization, Data Augmentation	ODCA-YOLO	mAP: 78.5%	The model's performance in practical industrial applications might be limited by its computational demand and the quality of input images, which must be high-resolution and well-prepped for accurate defect detection.
Keshun *et al.* [134]	Custom Dataset: Dataset I & Dataset II	Denoising, Image Enhancement	Semantic Segmentation using HALCON	mAP: Dataset I: 0.893, Dataset II: 0.905	The DL model can extract coordinate information of concentrate zoning but cannot accurately predict concentrate grade and recovery rate due to the lack of obtained zoning image features.

- Augmented Reality (AR) and Virtual Reality (VR): DL and CV technologies are enhancing AR and VR applications in manufacturing, enabling immersive experiences for training, maintenance, and product visualization. AR overlays real-time information onto the physical environment, providing operators contextual insights and guidance. VR simulations allow for realistic training scenarios and virtual prototyping, reducing time-to-market and improving product design.
- Cloud Computing and Edge Computing: DL and CV algorithms are deployed in cloud-based and edge computing environments. Cloud computing offers vast computational resources to train complex DL models and perform large-scale data analytics. On the other hand, edge computing brings processing closer to the data source, enabling real-time decision-making and reducing latency in critical applications like autonomous vehicles and remote monitoring systems.
- Blockchain Technology: Blockchain technology is increasingly being explored to enhance data security, traceability, and transparency in manufacturing supply chains. DL and CV techniques can be leveraged to analyze blockchain data for anomaly detection, fraud prevention, and quality assurance across the entire product lifecycle.

Integrating DL and CV technologies with other Industry 4.0 components drives a paradigm shift in industrial product management, paving the way for smarter, more agile, and interconnected manufacturing ecosystems. As these technologies continue to mature and evolve, their collective impact on efficiency, quality, and innovation in manufacturing will become even more pronounced.

VII. CURRENT LIMITATIONS AND FUTURE WORKS

While integrating Deep Learning and Computer Vision technologies into industrial product management holds immense promise, several challenges and limitations must be addressed to realize their full potential. This section delves into the complexities and barriers hindering these advanced technologies' seamless adoption and implementation in industrial settings.

A. LIMITATIONS OF CURRENT RESEARCH

Table 9 presents a comprehensive overview of the limitations encountered in current research on Deep Learning and Computer Vision technologies in industrial product

TABLE 9. Limitations of current research and potential solutions.

Types of Limitations of Current Research	Explanation of Limitation	Potential Solutions
Data Quality and Quantity	Limited availability of high-quality labelled datasets hamper the training of robust DL and CV models.	Collecting more labelled data through crowdsourcing, active learning, or data augmentation techniques.
Model Interpretability	Lack of interpretability in DL models impede understanding of decision-making processes, hindering trust and acceptance.	Developing explainable AI techniques such as model visualization, feature importance analysis, and attention mechanisms.
Robustness to Environmental Variability	DL and CV models may struggle to generalize to new environments due to variations in lighting, occlusions, and structural complexities.	Enhancing model robustness through domain adaptation, transfer learning, and robust optimization techniques.
Computational Resources	Training and deploying DL models require significant computational resources, posing challenges for resource-constrained environments.	Optimizing model architectures, implementing efficient algorithms, and leveraging cloud or distributed computing resources.
Overfitting and Generalization	DL models are susceptible to overfitting, resulting in poor performance when applied to unseen data due to their inability to generalize effectively.	Regularization techniques include dropout, L1/L2 regularization, early stopping, and cross-validation for hyperparameter tuning.
Labelling Bias and Imbalance	Biases and imbalances in labeled datasets can skew model predictions and compromise performance.	Employing bias correction methods, data balancing techniques, and incorporating fairness-aware learning approaches.
Ethical and Privacy Concerns	DL and CV technologies raise ethical concerns about privacy, bias, and discrimination, necessitating careful consideration and mitigation strategies.	Implementing privacy-preserving techniques, ensuring transparency and accountability in model development, and adhering to ethical guidelines and regulations.
Transferability of Models	DL models trained on one domain may not transfer well to other domains, limiting their applicability in diverse industrial settings.	Exploring domain adaptation methods, multi-task learning approaches, and model distillation techniques to improve transfer learning capabilities.
Model Degradation over Time	DL models may degrade in performance over time as data distributions shift or the model becomes outdated relative to evolving industrial processes.	Implementing continuous learning strategies, model retraining on updated data, and monitoring model drift to detect performance degradation.
Scalability and Efficiency	Scaling DL and CV solutions to handle large volumes of data in real-time can be challenging, especially in industrial settings with stringent latency requirements.	Utilizing parallel processing, model pruning, and hardware acceleration techniques to improve scalability and efficiency.

management, along with potential solutions to address these challenges. Each limitation, such as data quality and quantity, model interpretability, and computational resource constraints, is elucidated, detailing its specific challenges. Subsequently, corresponding potential solutions are proposed to mitigate these limitations, ranging from data augmentation techniques and model regularization to privacy-preserving methods and continuous learning strategies. The table offers valuable insights into the complexities of implementing DL and CV technologies in industrial settings by systematically delineating the challenges and potential remedies. It provides actionable approaches to overcome these barriers.

B. BARRIERS TO ADOPTION AND IMPLEMENTATION IN THE INDUSTRIAL CONTEXT

Table 10 provides an overview of the barriers hindering the adoption and implementation of Deep Learning and Computer Vision technologies in industrial contexts and potential solutions to overcome these hurdles. Each barrier, such as cost of implementation, lack of expertise, and regulatory compliance, is thoroughly explained, highlighting its specific challenges. Potential solutions are provided, offering actionable approaches to address these barriers effectively. By systematically delineating the challenges and potential remedies, the table offers valuable insights into the complexities of integrating DL and CV technologies into industrial settings. It provides practical strategies to facilitate their adoption and implementation.

C. FUTURE WORKS

Addressing the above mentioned challenges requires targeted research and development efforts. The following are potential future research directions to overcome the current limitations:

1) ADVANCED ILLUMINATION TECHNIQUES

Research into adaptive illumination techniques and robust pre-processing methods can help mitigate the impact of varying lighting conditions. Techniques such as photometric stereo and HDR imaging can be explored to enhance defect detection under different lighting scenarios. Developing models that can adapt to changing lighting conditions in real-time can also improve detection accuracy.

2) GENERALIZATION ACROSS DEFECT TYPES

Developing more robust models that can generalize across a broader range of defect types is essential. This can be achieved through the use of advanced architectures like transformers and the incorporation of domain adaptation

TABLE 10. Barriers to adoption and implementation in the industrial context and potential solutions.

Types of Barriers to Adoption and Implementation in the Industrial Context	Explanation of Barrier	Potential Solutions
Cost of Implementation	High upfront costs associated with infrastructure, software development, and training can deter organizations from adopting DL and CV technologies.	Exploring open-source solutions, cloud-based services, and cost-sharing models to reduce initial investment.
Lack of Expertise	Shortage of skilled professionals with expertise in computer vision, machine learning, and domain-specific knowledge impedes successful implementation.	Investing in employee training programs, partnering with educational institutions, and leveraging external expertise through consulting services.
Regulatory Compliance	Meeting regulatory requirements and ensuring compliance with industry standards adds complexity and time to the adoption process.	Collaborating with regulatory bodies, implementing robust compliance frameworks, and leveraging industry best practices to streamline regulatory compliance.
Legacy Systems Integration	Integrating DL and CV technologies with legacy systems poses compatibility and interoperability challenges.	Adopting modular architecture designs, utilizing middleware solutions, and implementing standardized communication protocols for seamless integration.
Risk Aversion and Inertia	Resistance to change and risk aversion within organizations may inhibit experimentation and innovation with new technologies.	Fostering a culture of innovation, incentivizing risk-taking behavior, and demonstrating the tangible benefits of DL and CV implementations through pilot projects.
Data Security Concerns	Concerns about data security and confidentiality may hinder the adoption of DL and CV solutions, particularly in sensitive industrial environments.	Implementing robust data encryption, access control mechanisms, and compliance with industry-specific security standards to mitigate data security risks.
The complexity of Deployment	Complexity involved in deploying and maintaining DL and CV systems, including software updates and troubleshooting, can be daunting for organizations.	Automating deployment processes, adopting DevOps practices, and utilizing containerization technologies for easier management and scalability.
Cultural Shifts	Cultural barriers, and organizational inertia may impede the cultural shift required to embrace digital transformation and innovation.	Facilitating open communication, promoting cross-functional collaboration, and fostering leadership buy-in to drive cultural change and adopt new technologies.
Vendor Lock-In	Dependency on specific vendors or proprietary technologies may limit flexibility and choice in adopting DL and CV solutions.	Embracing open standards, investing in interoperable solutions, and negotiating flexible vendor contracts to mitigate vendor lock-in risks.
Return on Investment Uncertainty	Uncertainty about the return on investment and long-term benefits of DL and CV implementations may discourage investment and commitment from stakeholders.	Conducting thorough cost-benefit analyses, defining clear performance metrics, and demonstrating tangible ROI through pilot projects and success stories to build stakeholder confidence and commitment.

techniques to improve model robustness. Techniques such as transfer learning and multi-task learning can also be employed to enhance the generalization capabilities of DL models.

3) INTEGRATION FRAMEWORKS

Creating standardized frameworks and APIs for the seamless integration of DL and CV technologies with existing manufacturing systems is crucial. Collaborative efforts between academia and industry can lead to the development of interoperable solutions that can be easily adopted by manufacturers. These frameworks should support modular integration, allowing for incremental adoption and scaling of DL and CV technologies.

4) SYNTHETIC DATA GENERATION

Leveraging synthetic data generation techniques can help address the issue of data scarcity. Techniques such as generative adversarial networks and simulation environments can be used to create large, annotated datasets for training DL models. Synthetic data can be designed to represent a wide range of defect types and conditions, providing valuable training data that may be difficult to obtain from real-world sources.

5) EFFICIENT MODEL ARCHITECTURES

Research into more efficient model architectures that require less computational power can make DL and CV technologies more accessible to SMEs. Techniques like model pruning, quantization, and the development of lightweight architectures can be explored to reduce the computational burden. Efficient models can also facilitate real-time processing and deployment on edge devices.

6) EDGE COMPUTING SOLUTIONS

Implementing edge computing solutions can help achieve real-time processing capabilities by performing inference at the edge of the network, closer to the data source. This approach can reduce latency and improve the responsiveness of defect detection systems. Edge computing can also enhance data privacy and security by minimizing the need to transmit sensitive data to central servers.

By addressing these challenges through focused research and development efforts, the integration of DL and CV technologies in manufacturing quality control can be significantly enhanced, leading to more efficient, accurate, and adaptable quality control processes.

VIII. CONCLUSION

This systematic review has comprehensively explored the significant advancements and persistent challenges associated with applying deep learning (DL) and computer vision (CV) techniques in industrial quality control, particularly for automated defect detection. While these technologies offer unparalleled potential in enhancing manufacturing processes, several challenges remain that hinder their widespread adoption and full integration. Key difficulties include the need for large, diverse datasets to train robust models, the complexity of deploying models in real-time production environments, and the variability in defect patterns that can lead to inconsistencies in detection accuracy. Additionally, issues such as the interpretability of DL models, the computational costs associated with real-time processing, and the need for seamless integration with existing manufacturing systems present significant hurdles.

Future research should prioritize the development of more generalized models capable of adapting to different manufacturing contexts with minimal retraining, the creation of more interpretable DL architectures, and the exploration of hybrid approaches that combine DL with traditional image processing techniques. Moreover, interdisciplinary collaboration between researchers, industry practitioners, and policymakers will be critical in addressing these challenges, ensuring that DL and CV technologies can be effectively and efficiently integrated into industrial workflows. By overcoming these barriers, the potential for DL and CV to revolutionize quality control and manufacturing processes can be fully realized, driving improvements in efficiency, product quality, and overall industry competitiveness.

ACKNOWLEDGMENT

The authors would like to thank the Advanced Machine Intelligence Research Laboratory—AMIR Laboratory for resources and supervision.

REFERENCES

[1] A. A. R. M. A. Ebayyeh and A. Mousavi, "A review and analysis of automatic optical inspection and quality monitoring methods in electronics industry," *IEEE Access*, vol. 8, pp. 183192–183271, 2020.

[2] T. Hussain, M. Hussain, H. Al-Aqrabi, T. Alsboui, and R. Hill, "A review on defect detection of electroluminescence-based photovoltaic cell surface images using computer vision," *Energies*, vol. 16, no. 10, p. 4012, May 2023.

[3] X. Zheng, S. Zheng, Y. Kong, and J. Chen, "Recent advances in surface defect inspection of industrial products using deep learning techniques," *Int. J. Adv. Manuf. Technol.*, vol. 113, nos. 1–2, pp. 35–58, Mar. 2021.

[4] L. Zhou, L. Zhang, and N. Konz, "Computer vision techniques in manufacturing," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 1, pp. 105–117, Jan. 2023.

[5] S. Dong, P. Wang, and K. Abbas, "A survey on deep learning and its applications," *Comput. Sci. Rev.*, vol. 40, May 2021, Art. no. 100379.

[6] R. Szeliski, *Computer Vision: Algorithms and Applications*. Cham, Switzerland: Springer, 2022, p. 925.

[7] S. Shahrabadi, Y. Castilla, M. Guevara, L. G. Magalhães, D. Gonzalez, and T. Adão, "Defect detection in the textile industry using image-based machine learning methods: A brief review," *J. Phys., Conf. Ser.*, vol. 2224, no. 1, Apr. 2022, Art. no. 012010.

[8] S. B. Jha and R. F. Babiceanu, "Deep CNN-based visual defect detection: Survey of current literature," *Comput. Ind.*, vol. 148, Jun. 2023, Art. no. 103911.

[9] H. Tercan and T. Meisen, "Machine learning and deep learning based predictive quality in manufacturing: A systematic review," *J. Intell. Manuf.*, vol. 33, no. 7, pp. 1879–1905, Oct. 2022.

[10] J. Yang, S. Li, Z. Wang, H. Dong, J. Wang, and S. Tang, "Using deep learning to detect defects in manufacturing: A comprehensive survey and current challenges," *Materials*, vol. 13, no. 24, p. 5755, Dec. 2020.

[11] V. Nasir and F. Sassani, "A review on deep learning in machining and tool monitoring: Methods, opportunities, and challenges," *Int. J. Adv. Manuf. Technol.*, vol. 115, nos. 9–10, pp. 2683–2709, Aug. 2021.

[12] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," *J. Manuf. Syst.*, vol. 48, pp. 144–156, Jul. 2018.

[13] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Mar. 2020.

[14] T. Wang and V. J. L. Gan, "Automated joint 3D reconstruction and visual inspection for buildings using computer vision and transfer learning," *Autom. Construct.*, vol. 149, May 2023, Art. no. 104810.

[15] L. Yu, E. Yang, C. Luo, and P. Ren, "AMCD: An accurate deep learning-based metallic corrosion detector for MAV-based real-time visual inspection," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 7, pp. 8087–8098, Jul. 2023.

[16] X. Li, W. Wang, L. Sun, B. Hu, L. Zhu, and J. Zhang, "Deep learning-based defects detection of certain aero-engine blades and vanes with DDSC-YOLOv5s," *Sci. Rep.*, vol. 12, no. 1, p. 13067, Jul. 2022.

[17] A. Upadhyay, J. Li, S. King, and S. Addepalli, "A deep-learning-based approach for aircraft engine defect detection," *Machines*, vol. 11, no. 2, p. 192, Feb. 2023.

[18] D. Wang, H. Xiao, and S. Huang, "Automatic defect recognition and localization for aeroengine turbine blades based on deep learning," *Aerospace*, vol. 10, no. 2, p. 178, Feb. 2023.

[19] Y. Keshun, W. Puzhou, and G. Yingkui, "Toward efficient and interpretative rolling bearing fault diagnosis via quadratic neural network with bi-LSTM," *IEEE Internet Things J.*, vol. 11, no. 13, pp. 23002–23019, Jul. 2024.

[20] Y. Keshun, Q. Guangqi, and G. Yingkui, "Optimizing prior distribution parameters for probabilistic prediction of remaining useful life using deep learning," *Rel. Eng. Syst. Saf.*, vol. 242, Feb. 2024, Art. no. 109793.

[21] C. El Hachem, G. Perrot, L. Painvin, J.-B. Ernst-Desmulier, and R. Couturier, "Welding seam classification in the automotive industry using deep learning algorithms," in *Proc. IEEE Int. Conf. Ind. 4.0, Artif. Intell., Commun. Technol. (IAICT)*, Jul. 2021, pp. 235–240.

[22] R. Miao, Z. Jiang, Q. Zhou, Y. Wu, Y. Gao, J. Zhang, and Z. Jiang, "Online inspection of narrow overlap weld quality using two-stage convolution neural network image recognition," *Mach. Vis. Appl.*, vol. 32, no. 1, pp. 1–14, Jan. 2021.

[23] J. Wang, C. Mu, S. Mu, R. Zhu, and H. Yu, "Welding seam detection and location: Deep learning network-based approach," *Int. J. Pressure Vessels Piping*, vol. 202, Apr. 2023, Art. no. 104893.

[24] Y. Keshun and L. Huizhong, "Feature detection of mineral zoning in spiral slope flow under complex conditions based on improved YOLOv5 algorithm," *Phys. Scripta*, vol. 99, no. 1, Jan. 2024, Art. no. 016001.

[25] K. You, G. Qiu, and Y. Gu, "An efficient lightweight neural network using BiLSTM-SCN-CBAM with PCA-ICEEMDAN for diagnosing rolling bearing faults," *Meas. Sci. Technol.*, vol. 34, no. 9, Sep. 2023, Art. no. 094001.

[26] Y. Keshun, Q. Guangqi, and G. Yingkui, "Remaining useful life prediction of lithium-ion batteries using EM-PF-SSA-SVR with gamma stochastic process," *Meas. Sci. Technol.*, vol. 35, no. 1, Jan. 2024, Art. no. 015015.

[27] J. Lian, L. Wang, T. Liu, X. Ding, and Z. Yu, "Automatic visual inspection for printed circuit board via novel mask R-CNN in smart city applications," *Sustain. Energy Technol. Assessments*, vol. 44, Apr. 2021, Art. no. 101032.

[28] J. Kim, J. Ko, H. Choi, and H. Kim, "Printed circuit board defect detection using deep learning via a skip-connected convolutional autoencoder," *Sensors*, vol. 21, no. 15, p. 4968, Jul. 2021.

[29] T. T. A. Pham, D. K. T. Thoi, H. Choi, and S. Park, "Defect detection in printed circuit boards using semi-supervised learning," *Sensors*, vol. 23, no. 6, p. 3246, Mar. 2023.

[30] J. Wang, P. Gao, J. Zhang, C. Lu, and B. Shen, "Knowledge augmented broad learning system for computer vision based mixed-type defect detection in semiconductor manufacturing," *Robot. Comput.-Integr. Manuf.*, vol. 81, Jun. 2023, Art. no. 102513.

[31] T.-Y. Kim, S. Park, C.-K. Lim, J.-H. Choi, C.-H. An, L.-H. Song, and D.-N. Kim, "Deep learning-based detection of defects in wafer buffer zone during semiconductor packaging process," *Multiscale Sci. Eng.*, vol. 6, no. 1, pp. 25–32, Mar. 2024.

[32] F. López de la Rosa, J. L. Gómez-Sirvent, R. Morales, R. Sánchez-Reolid, and A. Fernández-Caballero, "Defect detection and classification on semiconductor wafers using two-stage geometric transformation-based data augmentation and SqueezeNet lightweight convolutional neural network," *Comput. Ind. Eng.*, vol. 183, Sep. 2023, Art. no. 109549.

[33] K. Limam, S. Cheema, S. Mouhoubi, and F. D. Freijedo, "Deep learning-based visual recognition for inline defects in production of semiconductors," *IEEE J. Emerg. Sel. Topics Ind. Electron.*, vol. 5, no. 1, pp. 203–211, Jan. 2024.

[34] Q. Zhang, K. Zhang, K. Pan, and W. Huang, "Image defect classification of surface mount technology welding based on the improved ResNet model," *J. Eng. Res.*, vol. 12, no. 2, pp. 154–162, Jun. 2024.

[35] H. Wu, R. Lei, and Y. Peng, "PCBNet: A lightweight convolutional neural network for defect inspection in surface mount technology," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–14, 2022.

[36] S. Dlamini, C.-F.-J. Kuo, and S.-M. Chao, "Developing a surface mount technology defect detection system for mounted devices on printed circuit boards using a MobileNetV2 with feature pyramid network," *Eng. Appl. Artif. Intell.*, vol. 121, May 2023, Art. no. 105875.

[37] J. Jing, Z. Wang, M. Rätsch, and H. Zhang, "Mobile-Unet: An efficient convolutional neural network for fabric defect detection," *Textile Res. J.*, vol. 92, nos. 1–2, pp. 30–42, Jan. 2022.

[38] G. Hu, J. Huang, Q. Wang, J. Li, Z. Xu, and X. Huang, "Unsupervised fabric defect detection based on a deep convolutional generative adversarial network," *Textile Res. J.*, vol. 90, nos. 3–4, pp. 247–270, Feb. 2020.

[39] L. Cheng, J. Yi, A. Chen, and Y. Zhang, "Fabric defect detection based on separate convolutional UNet," *Multimedia Tools Appl.*, vol. 82, no. 2, pp. 3101–3122, Jan. 2023.

[40] Y. Zhao, K. Hao, H. He, X. Tang, and B. Wei, "A visual long-short-term memory based integrated CNN model for fabric defect image classification," *Neurocomputing*, vol. 380, pp. 259–270, Mar. 2020.

[41] M. A. Iqbal Hussain, B. Khan, Z. Wang, and S. Ding, "Woven fabric pattern recognition and classification based on deep convolutional neural networks," *Electronics*, vol. 9, no. 6, p. 1048, Jun. 2020.

[42] H.-W. Zhang, Q.-L. Tan, S. Lu, Z.-Q. Ge, and D. Gu, "Yarn-dyed fabric defect detection using U-shaped de-noising convolutional auto-encoder," in *Proc. IEEE 9th Data Driven Control Learn. Syst. Conf. (DDCLS)*, Nov. 2020, pp. 18–24.

[43] J. Xiao, A. K. Aggarwal, U. K. Rage, V. Katiyar, and R. Avtar, "Deep learning-based spatiotemporal fusion of unmanned aerial vehicle and satellite reflectance images for crop monitoring," *IEEE Access*, vol. 11, pp. 85600–85614, 2023.

[44] H. R. Khan, Z. Gillani, M. H. Jamal, A. Athar, M. T. Chaudhry, H. Chao, Y. He, and M. Chen, "Early identification of crop type for smallholder farming systems using deep learning on time-series Sentinel-2 imagery," *Sensors*, vol. 23, no. 4, p. 1779, Feb. 2023.

[45] I. Blekanov, A. Molin, D. Zhang, E. Mitrofanov, O. Mitrofanova, and Y. Li, "Monitoring of grain crops nitrogen status from uav multispectral images coupled with deep learning approaches," *Comput. Electron. Agricult.*, vol. 212, Sep. 2023, Art. no. 108047.

[46] D. Cama-Pinto, M. Damas, J. A. Holgado-Terriza, F. M. Arrabal-Campos, J. A. Martínez-Lao, A. Cama-Pinto, and F. Manzano-Agugliaro, "A deep learning model of radio wave propagation for precision agriculture and sensor system in greenhouses," *Agronomy*, vol. 13, no. 1, p. 244, Jan. 2023.

[47] P. Kumar, A. Udayakumar, A. A. Kumar, K. S. Kannan, and N. Krishnan, "Multiparameter optimization system with DCNN in precision agriculture for advanced irrigation planning and scheduling based on soil moisture estimation," *Environ. Monitor. Assessment*, vol. 195, no. 1, p. 13, Jan. 2023.

[48] D. Banerjee, V. Kukreja, S. Hariharan, and V. Sharma, "Precision agriculture: Classifying banana leaf diseases with hybrid deep learning models," in *Proc. IEEE 8th Int. Conf. Converg. Technol. (I2CT)*, Apr. 2023, pp. 1–5.

[49] S. I. Moazzam, U. S. Khan, W. S. Qureshi, T. Nawaz, and F. Kunwar, "Towards automated weed detection through two-stage semantic segmentation of tobacco and weed pixels in aerial imagery," *Smart Agricult. Technol.*, vol. 4, Aug. 2023, Art. no. 100142.

[50] P. Ong, K. S. Teo, and C. K. Sia, "UAV-based weed detection in Chinese cabbage using deep learning," *Smart Agricult. Technol.*, vol. 4, Aug. 2023, Art. no. 100181.

[51] S. Tummapudi, S. S. Sadhu, S. N. Simhadri, S. N. T. Damarla, and M. Bhukya, "Deep learning based weed detection and elimination in agriculture," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, Apr. 2023, pp. 147–151.

[52] G. Morales, J. W. Sheppard, P. B. Hegedus, and B. D. Maxwell, "Improved yield prediction of winter wheat using a novel two-dimensional deep regression neural network trained via remote sensing," *Sensors*, vol. 23, no. 1, p. 489, Jan. 2023.

[53] R. Tanabe, T. Matsui, and T. S. T. Tanaka, "Winter wheat yield prediction using convolutional neural networks and UAV-based multispectral imagery," *Field Crops Res.*, vol. 291, Feb. 2023, Art. no. 108786.

[54] N. Gururaj, V. Vinod, and K. Vijayakumar, "Deep grading of mangoes using convolutional neural network and computer vision," *Multimedia Tools Appl.*, vol. 82, no. 25, pp. 39525–39550, Oct. 2023.

[55] S. Dharmadhikari and A. Basak, "Fatigue damage detection of aerospace-grade aluminum alloys using feature-based and feature-less deep neural networks," *Mach. Learn. with Appl.*, vol. 7, Mar. 2022, Art. no. 100247.

[56] M. Le and J. Lee, "Ultrasonic testing of corrosion in aircraft rivet using spiking neural network," *J. Nondestruct. Eval.*, vol. 42, no. 3, p. 78, Sep. 2023.

[57] H. T. Quan, D. D. Huy, N. T. Hoan, and N. T. Duc, "Deep learning-based automatic detection of defective tablets in pharmaceutical manufacturing," in *Proc. 8th Int. Conf. Develop. Biomed. Eng. Vietnam*, vol. 85. Cham, Switzerland: Springer, 2020, pp. 789–801.

[58] M. Ficzere, L. A. Mészáros, N. Kállai-Szabó, A. Kovács, I. Antal, Z. K. Nagy, and D. L. Galata, "Real-time coating thickness measurement and defect recognition of film coated tablets with machine vision and deep learning," *Int. J. Pharmaceutics*, vol. 623, Jul. 2022, Art. no. 121957.

[59] K. C. Laxman, N. Tabassum, L. Ai, C. Cole, and P. Ziehl, "Automated crack detection and crack depth prediction for reinforced concrete structures using deep learning," *Construct. Building Mater.*, vol. 370, Mar. 2023, Art. no. 130709.

[60] D. Tabernik, M. Šuc, and D. Skočaj, "Automated detection and segmentation of cracks in concrete surfaces using joined segmentation and classification deep neural network," *Construct. Building Mater.*, vol. 408, Dec. 2023, Art. no. 133582.

[61] Z. Li, X. Wei, M. Hassaballah, Y. Li, and X. Jiang, "A deep learning model for steel surface defect detection," *Complex Intell. Syst.*, vol. 10, no. 1, pp. 885–897, Feb. 2024.

[62] K. Demir, M. Ay, M. Cavas, and F. Demir, "Automated steel surface defect detection and classification using a new deep learning-based approach," *Neural Comput. Appl.*, vol. 35, no. 11, pp. 8389–8406, Apr. 2023.

[63] W.-H. Lim, M. B. Bonab, and K. H. Chua, "An aggressively pruned CNN model with visual attention for near real-time wood defects detection on embedded processors," *IEEE Access*, vol. 11, pp. 36834–36848, 2023.

[64] W. Cui, Z. Li, A. Duanmu, S. Xue, Y. Guo, C. Ni, T. Zhu, and Y. Zhang, "CCG-YOLOv7: A wood defect detection model for small targets using improved YOLOv7," *IEEE Access*, vol. 12, pp. 10575–10585, 2024.

[65] D. Rački, D. Tomaževič, and D. Skočaj, "Detection of surface defects on pharmaceutical solid oral dosage forms with convolutional neural networks," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 631–650, Jan. 2022.

[66] W. Ouyang, B. Xu, J. Hou, and X. Yuan, "Fabric defect detection using activation layer embedded convolutional neural network," *IEEE Access*, vol. 7, pp. 70130–70140, 2019.

[67] O. O. Abayomi-Alli, R. Damaševič ius, S. Misra, and R. Maskeliūnas, "Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning," *Expert Syst.*, vol. 38, no. 7, p. 12746, Nov. 2021.

[68] S. Ashok, G. Kishore, V. Rajesh, S. Suchitra, S. G. G. Sophia, and B. Pavithra, "Tomato leaf disease detection using deep learning techniques," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2020, pp. 979–983.

[69] O. Asif, S. A. Haider, S. R. Naqvi, J. F. W. Zaki, K.-S. Kwak, and S. M. R. Islam, "A deep learning model for remaining useful life prediction of aircraft turbofan engine on C-MAPSS dataset," *IEEE Access*, vol. 10, pp. 95425–95440, 2022.

[70] S. Alsubai, A. K. Dutta, A. H. Alkhayyat, M. M. Jaber, A. H. Abbas, and A. Kumar, "Hybrid deep learning with improved salp swarm optimization based multi-class grape disease classification model," *Comput. Electr. Eng.*, vol. 108, May 2023, Art. no. 108733.

[71] I.-C. Chen, R.-C. Hwang, and H.-C. Huang, "PCB defect detection based on deep learning algorithm," *Processes*, vol. 11, no. 3, p. 775, Mar. 2023.

[72] M. M. Khodier, S. M. Ahmed, and M. S. Sayed, "Complex pattern jacquard fabrics defect detection using convolutional neural networks and multispectral imaging," *IEEE Access*, vol. 10, pp. 10653–10660, 2022.

[73] Z. U. Rehman, M. A. Khan, F. Ahmed, R. Damaševič ius, S. R. Naqvi, W. Nisar, and K. Javed, "Recognizing apple leaf diseases using a novel parallel real-time processing framework based on MASK RCNN and transfer learning: An application for smart agriculture," *IET Image Process.*, vol. 15, no. 10, pp. 2157–2168, Aug. 2021.

[74] F. T. J. Faria, M. B. Moin, A. A. Wase, M. R. Sani, K. M. Hasib, and M. S. Alam, "Classification of potato disease with digital image processing technique: A hybrid deep learning framework," in *Proc. IEEE 13th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Mar. 2023, pp. 820–826.

[75] L. O. Chua, *CNN: A Paradigm for Complexity*, vol. 31. Singapore: World Scientific, 1998, pp. 529–837.

[76] G. Latif, S. E. Abdelhamid, R. E. Mallouhy, J. Alghazo, and Z. A. Kazimi, "Deep learning utilization in agriculture: Detection of Rice plant diseases using an improved CNN model," *Plants*, vol. 11, no. 17, p. 2230, Aug. 2022.

[77] J.-C. Chien, M.-T. Wu, and J.-D. Lee, "Inspection and classification of semiconductor wafer surface defects using CNN deep learning networks," *Appl. Sci.*, vol. 10, no. 15, p. 5340, Aug. 2020.

[78] J. Huang, H. He, R. Lv, G. Zhang, Z. Zhou, and X. Wang, "Non-destructive detection and classification of textile fibres based on hyperspectral imaging and 1D-CNN," *Analytica Chim. Acta*, vol. 1224, Sep. 2022, Art. no. 340238.

[79] Y. Keshun, Q. Guangqi, and G. Yingkui, "A 3-D attention-enhanced hybrid neural network for turbofan engine remaining life prediction using CNN and BiLSTM models," *IEEE Sensors J.*, vol. 24, no. 14, pp. 21893–21905, Jul. 2024.

[80] A. Khaouane, L. Khaouane, S. Ferhat, and S. Hanini, "Deep learning for drug development: Using CNNs in MIA-QSAR to predict plasma protein binding of drugs," *AAPS PharmSciTech*, vol. 24, no. 8, p. 232, Nov. 2023.

[81] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[82] N. Hnoohom, N. Maitrichit, P. Chotivatunyu, V. Sornlertlamvanich, S. Mekruksavanich, and A. Jitpattanakul, "Blister package classification using ResNet-101 for identification of medication," in *Proc. 25th Int. Comput. Sci. Eng. Conf. (ICSEC)*, Nov. 2021, pp. 406–410.

[83] P. Tchatchoua, G. Graton, M. Ouladsine, and J.-F. Christaud, "Application of 1D ResNet for multivariate fault detection on semiconductor manufacturing equipment," *Sensors*, vol. 23, no. 22, p. 9099, Nov. 2023.

[84] A. Stephen, A. Punitha, and A. Chandrasekar, "Designing self attention-based ResNet architecture for Rice leaf disease classification," *Neural Comput. Appl.*, vol. 35, no. 9, pp. 6737–6751, Mar. 2023.

[85] H. Zhang, L. Jiang, and C. Li, "CS-ResNet: Cost-sensitive residual convolutional neural network for PCB cosmetic defect detection," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115673.

[86] R. Liu, Z. Yu, Q. Fan, Q. Sun, and Z. Jiang, "The improved method in fabric image classification using convolutional neural network," *Multimedia Tools Appl.*, vol. 83, no. 3, pp. 6909–6924, Jan. 2024.

[87] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[88] S. Kulkarni, H. N. Shravan, M. V. Veeresh, P. D. Shenoy, and K. R. Venugopal, "Detection of coffee leaf diseases using DenseNet-264 and SLIC segmentation," in *Proc. 10th IEEE Uttar Pradesh Sect. Int. Conf. Electr., Electron. Comput. Eng. (UPCON)*, Dec. 2023, pp. 769–774.

[89] G. Wang, Y. Zhang, F. Zhang, and Z. Wu, "An ensemble method with DenseNet and evidential reasoning rule for machinery fault diagnosis under imbalanced condition," *Measurement*, vol. 214, Jun. 2023, Art. no. 112806.

[90] Z. Zhu, G. Han, G. Jia, and L. Shu, "Modified DenseNet for automatic fabric defect detection with edge computing for minimizing latency," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9623–9636, Oct. 2020.

[91] L. R. Medsker and L. Jain, "Recurrent neural networks," *Design Appl.*, vol. 5, nos. 64–67, p. 2, 2001.

[92] Z. Hu, Q. Zhao, and J. Wang, "The prediction model of cotton yarn quality based on artificial recurrent neural network," in *Proc. Int. Conf. Appl. Techn. Cyber Intell. (ATCI)*, vol. 1017. Cham, Switzerland: Springer, 2019, pp. 857–866.

[93] P. Kumari, V. Goswami, and R. S. Pundir, "Recurrent neural network architecture for forecasting banana prices in Gujarat, India," *PLoS ONE*, vol. 18, no. 6, Jun. 2023, Art. no. e0275702.

[94] P. R. Puentes, N. Valderrama, C. González, L. Daza, C. Muñoz-Camargo, J. C. Cruz, and P. Arbeláez, "PharmaNet: Pharmaceutical discovery with deep recurrent neural networks," *PLoS ONE*, vol. 16, no. 4, Apr. 2021, Art. no. e0241728.

[95] Y. Wang, M. Perry, D. Whitlock, and J. W. Sutherland, "Detecting anomalies in time series data from a manufacturing system using recurrent neural networks," *J. Manuf. Syst.*, vol. 62, pp. 823–834, Jan. 2022.

[96] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[97] K. S. Kumar and M. R. Bai, "LSTM based texture classification and defect detection in a fabric," *Meas., Sensors*, vol. 26, Apr. 2023, Art. no. 100603.

[98] W. H. Chung, Y. H. Gu, and S. J. Yoo, "CHP engine anomaly detection based on parallel CNN-LSTM with residual blocks and attention," *Sensors*, vol. 23, no. 21, p. 8746, Oct. 2023.

[99] S. Ray, A. Lama, P. Mishra, T. Biswas, S. Sankar Das, and B. Gurung, "An ARIMA-LSTM model for predicting volatile agricultural price series with random forest technique," *Appl. Soft Comput.*, vol. 149, Dec. 2023, Art. no. 110939.

[100] J. Qian, X. Qiu, X. Tan, Q. Li, J. Chen, and X. Jiang, "An attentive LSTM based approach for adverse drug reactions prediction," *Appl. Intell.*, vol. 53, no. 5, pp. 4875–4889, Jun. 2022.

[101] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Jun. 2009.

[102] B. Chellapandi, M. Vijayalakshmi, and S. Chopra, "Comparison of pre-trained models using transfer learning for detecting plant disease," in *Proc. Int. Conf. Comput., Commun., Intell. Syst. (ICCCIS)*, Feb. 2021, pp. 383–387.

[103] Y. Zhu, R. Liu, G. Hu, X. Chen, and W. Li, "Accurate identification of cashmere and wool fibers based on enhanced ShuffleNetV2 and transfer learning," *J. Big Data*, vol. 10, no. 1, p. 152, Oct. 2023.

[104] S. Katsigiannis, S. Seyedzadeh, A. Agapiou, and N. Ramzan, "Deep learning for crack detection on masonry façades using limited data and transfer learning," *J. Building Eng.*, vol. 76, Oct. 2023, Art. no. 107105.

[105] H. Iwata, Y. Hayashi, A. Hasegawa, K. Terayama, and Y. Okuno, "Classification of scanning electron microscope images of pharmaceutical excipients using deep convolutional neural networks with transfer learning," *Int. J. Pharmaceutics, X*, vol. 4, Dec. 2022, Art. no. 100135.

[106] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 2672–2680.

[107] C. Li, W. Xia, and Z. Jiang, "Weak feature defect generation with GAN for faster RCNN based PCB defect detection," in *Proc. 8th Int. Conf. Data Sci. Cyberspace (DSC)*, Aug. 2023, pp. 306–312.

[108] S. Lamba, P. Saini, J. Kaur, and V. Kukreja, "Optimized classification model for plant diseases using generative adversarial networks," *Innov. Syst. Softw. Eng.*, vol. 19, no. 1, pp. 103–115, Mar. 2023.

[109] B. Xu and C. Liu, "Pavement crack detection algorithm based on generative adversarial network and convolutional neural network under small samples," *Measurement*, vol. 196, Jun. 2022, Art. no. 111219.

[110] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.

[111] A. Ramlatchan and Y. Li, "Image synthesis using conditional GANs for selective laser melting additive manufacturing," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2022, pp. 1–8.

[112] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2006, pp. 1735–1742.

[113] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[114] J. K. Chow, Z. Su, J. Wu, P. S. Tan, X. Mao, and Y. H. Wang, "Anomaly detection of defects on concrete structures with the convolutional autoencoder," *Adv. Eng. Informat.*, vol. 45, Aug. 2020, Art. no. 101105.

[115] C. Qu, Z. Zhou, Z. Liu, and S. Jia, "Predictive anomaly detection for marine diesel engine based on echo state network and autoencoder," *Energy Rep.*, vol. 8, pp. 998–1003, Jul. 2022.

[116] E. K. Yılmaz, K. Adem, S. Kılıçarslan, and H. A. Aydın, "Classification of lemon quality using hybrid model based on stacked AutoEncoder and convolutional neural network," *Eur. Food Res. Technol.*, vol. 249, no. 6, pp. 1655–1667, Jun. 2023.

[117] M. H. Zubayer, C. Zhang, W. Liu, Y. Wang, and H. M. Imdadul, "Automatic defect detection of jet engine turbine and compressor blade surface coatings using a deep learning-based algorithm," *Coatings*, vol. 14, no. 4, p. 501, Apr. 2024.

[118] Y. Ma, M. Liu, Y. Zhang, X. Wang, and Y. Wang, "SPDP-net: A semantic prior guided defect perception network for automated aero-engine blades surface visual inspection," *IEEE Trans. Autom. Sci. Eng.*, early access, Apr. 2, 2024, doi: 10.1109/TASE.2024.3383451.

[119] G. U. Kaya, "Development of hybrid optical sensor based on deep learning to detect and classify the micro-size defects in printed circuit board," *Measurement*, vol. 206, Jan. 2023, Art. no. 112247.

[120] Y. Zhang, T. Han, B. Wei, K. Hao, and L. Gao, "A spatial–spectral adaptive learning model for textile defect images recognition with few labeled data," *Complex Intell. Syst.*, vol. 9, no. 6, pp. 6359–6371, Dec. 2023.

[121] N. Alruwais, E. Alabdulkreem, K. Mahmood, R. Marzouk, M. Assiri, A. A. Abdelmageed, S. Abdelbagi, and S. Drar, "Hybrid mutation moth flame optimization with deep learning-based smart fabric defect detection," *Comput. Electr. Eng.*, vol. 108, May 2023, Art. no. 108706.

[122] G. Revathy and R. Kalaivani, "Fabric defect detection and classification via deep learning-based improved mask RCNN," *Signal, Image Video Process.*, vol. 18, no. 3, pp. 2183–2193, Apr. 2024.

[123] M. T. Mallick, S. Biswas, A. K. Das, H. N. Saha, A. Chakrabarti, and N. Deb, "Deep learning based automated disease detection and pest classification in Indian Mung bean," *Multimedia Tools Appl.*, vol. 82, no. 8, pp. 12017–12041, Mar. 2023.

[124] W. Yang, P. Shen, Z. Ye, Z. Zhu, C. Xu, Y. Liu, and L. Mei, "Adversarial training collaborating multi-path context feature aggregation network for maize disease density prediction," *Processes*, vol. 11, no. 4, p. 1132, Apr. 2023.

[125] A. Dash, P. K. Sethy, and S. K. Behera, "Maize disease identification based on optimized support vector machine using deep feature of DenseNet201," *J. Agricult. Food Res.*, vol. 14, Dec. 2023, Art. no. 100824.

[126] K. Roy, S. S. Chaudhuri, J. Frnda, S. Bandopadhyay, I. J. Ray, S. Banerjee, and J. Nedoma, "Detection of tomato leaf diseases for agro-based industries using novel PCA DeepNet," *IEEE Access*, vol. 11, pp. 14983–15001, 2023.

[127] X. Zhang, Z. Wu, C. Cao, K. Luo, K. Qin, Y. Huang, and J. Cao, "Design and operation of a deep-learning-based fresh tea-leaf sorting robot," *Comput. Electron. Agricult.*, vol. 206, Mar. 2023, Art. no. 107664.

[128] J. Padmapriya and T. Sasilatha, "Deep learning based multi-labelled soil classification and empirical estimation toward sustainable agriculture," *Eng. Appl. Artif. Intell.*, vol. 119, Mar. 2023, Art. no. 105690.

[129] E. Li, L. Wang, Q. Xie, R. Gao, Z. Su, and Y. Li, "A novel deep learning method for maize disease identification based on small sample-size and complex background datasets," *Ecol. Informat.*, vol. 75, Jul. 2023, Art. no. 102011.

[130] F. Özgenel, 2018, "Concrete crack images for classification," *Mendeley Data*, doi: 10.17632/5y9wdsg2zt.2.

[131] H. Li, Z. Yue, J. Liu, Y. Wang, H. Cai, K. Cui, and X. Chen, "SCCDNet: A pixel-level crack segmentation network," *Appl. Sci.*, vol. 11, no. 11, p. 5074, May 2021.

[132] Y. Tu, Z. Ling, S. Guo, and H. Wen, "An accurate and real-time surface defects detection method for sawn lumber," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[133] R. Wang, F. Liang, B. Wang, and X. Mou, "ODCA-YOLO: An omni-dynamic convolution coordinate attention-based Yolo for wood defect detection," *Forests*, vol. 14, no. 9, p. 1885, Sep. 2023.

[134] Y. Keshun, W. Chengyu, and L. Huizhong, "Research on intelligent implementation of the beneficiation process of shaking table," *Minerals Eng.*, vol. 199, Aug. 2023, Art. no. 108108.

MD RAISUL ISLAM received the B.Sc. degree in electrical and electronic engineering from Bangladesh University of Business and Technology, Bangladesh. He is currently pursuing the master's degree in manufacturing engineering with the School of Engineering and Technology, Western Illinois University, USA. His research interests include the application of artificial intelligence in industrial automation and manufacturing engineering.

MD ZAKIR HOSSAIN ZAMIL received the Bachelor of Science degree in computer science and engineering from Bangladesh University of Business and Technology, Bangladesh. He is currently pursuing the master's degree in computer science and engineering with the School of Engineering and Technology, Western Illinois University, USA. His research interests include artificial intelligence (AI), deep learning (DL), machine learning (ML), and computer vision (CV).

MD ESHMAM RAYED received the B.Sc. degree in computer science and engineering (CSE) with a major in software engineering from American International University-Bangladesh (AIUB), in 2023. He finished the internship as a Research Intern with the Advanced Machine Intelligence Research Laboratory (AMIR Laboratory). He is familiar with various programming languages, including C/C++, Python, TensorFlow, and PyTorch. His research interests include deep learning, machine learning, human activity recognition, medical imaging, federated learning, and natural language processing.

MD MOHSIN KABIR received the Bachelor of Science degree in computer science and engineering from Bangladesh University of Business and Technology (BUBT), in 2021. He completed Erasmus Mundus Joint Master's in Intelligent Field Robotics Systems from the University of Girona, Spain, and End University, Hungary, in 2024. He was a Lecturer with the Department of Computer Science and Engineering, BUBT, where he has a rich background in research, having held positions, such as Research Assistant and a Lead Researcher with the Advanced Machine Intelligence Research Laboratory. He has made significant contributions to several top-tier research laboratories globally, including the Database System Laboratory, The University of Aizu, Japan. His research work is extensively published, with over 30 articles in high-impact journals such as *Scientific Reports*, *Computers and Electrical Engineering*, *Informatics in Medicine Unlocked*, *Natural Language Processing Journal*, *Heliyon*, *Computers & Security*, *Healthcare Analytics*, *Journal of Agriculture and Food Research*, IEEE ACCESS, *Sensors*, *PLOS One*, *Cognitive Computation and Systems*, *International Journal of Information Management Data Insights*, and *Mathematics* Additionally, his contributions extend to over 12 conference papers and three book chapters in machine learning and artificial intelligence. His research interests include broad and deep, focusing on artificial intelligence, machine learning, deep learning, computer vision, the Internet of Things (IoT), and robotics.

M. F. MRIDHA (Senior Member, IEEE) received the Ph.D. degree in AI/ML from Jahangirnagar University, in 2017. He is currently an Associate Professor with the Department of Computer Science, American International University-Bangladesh (AIUB). Before that, he was an Associate Professor and the Chairperson of the Department of CSE, Bangladesh University of Business and Technology. He was the CSE Department Faculty Member of the University of Asia Pacific and as the Graduate Head, from 2012 to 2019. His research experience, within both academia and industry, results in over 120 journals and conference publications. His research work contributed to the reputed journal of *Scientific Reports* (Nature), *Knowledge-Based Systems*, *Artificial Intelligence Review*, IEEE ACCESS, *Sensors*, *Cancers*, and *Applied Sciences*. For more than ten years, he has been with the masters and undergraduate students as a supervisor of their thesis work. His research interests include artificial intelligence (AI), machine learning, deep learning, natural language processing (NLP), and big data analysis. He has served as a program committee member for several international conferences/workshops. He served as an Associate Editor for several journals, including *PLOS One* journal. He has served as a Reviewer for reputed journals and IEEE conferences like HONET, ICIEV, ICCIT, IJCCI, ICAEE, ICCAIE, ICSIPA, SCORED, ISIEA, APACE, ICOS, ISCAIE, BEIAC, ISWTA, IC3e, ISWTA, CoAST, icIVPR, ICSCT, 3ICT, and DATA21.

SATOSHI NISHIMURA (Member, IEEE) received the B.E. degree from Tohoku University in 1987, and the M.Sc. and D.Sc. degrees in information science from The University of Tokyo, in 1989 and 1995, respectively. He is currently a Senior Associate Professor with The University of Aizu. His research interests include computer graphics and computer music. He is a member of ACM and IPSJ.

JUNGPIL SHIN (Senior Member, IEEE) received the B.Sc. degree in computer science and statistics and the M.Sc. degree in computer science from Pusan National University, South Korea, in 1990 and 1994, respectively, and the Ph.D. in computer science and communication engineering from Kyushu University, Japan, in 1999, under a scholarship from Japanese government (MEXT). He was an Associate Professor, a Senior Associate Professor, and a Full Professor with the School of Computer Science and Engineering, The University of Aizu, Japan, in 1999, 2004, and 2019, respectively. He has co-authored more than 400 published papers for widely cited journals and conferences. His research interests include pattern recognition, image processing, computer vision, machine learning, human–computer interaction, nontouch interfaces, human gesture recognition, automatic control, Parkinson's disease diagnosis, ADHD diagnosis, user authentication, machine intelligence, bioinformatics, handwriting analysis, recognition, and synthesis. He is a member of ACM, IEICE, IPSJ, KISS, and KIPS. He served as the program chair and as a program committee member for numerous international conferences. He serves as an Editor for IEEE journals, Springer, Sage, Taylor & Francis, *Sensors* (MDPI) and *Electronics* (MDPI), and Tech Science. He serves as an Editorial Board Member for *Scientific Reports*. He serves as a reviewer for several major IEEE and SCI journals.