YOLO-BASED VISUAL INSPECTION IN INDUSTRY 4.0: A COMPREHENSIVE REVIEW OF SMART MANUFACTURING APPLICATIONS

C. Dewi[1*], E.K. Pradibta Fury[1], J. S. Rosario Putra[1], E.J. Feodora Aritonang[1], A.B. Setia Permana[1], M.H.F. Md Fauadi[2], S. Aprius Sutresno[3], and D. Riantama[3]

[1]Department of Information Technology, Satya Wacana Christian University, Jl. Diponegoro No. 52-60, Salatiga, 50711, Indonesia.

[2]Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka, 76100 Hang Tuah Jaya, Melaka, Malaysia.

[3]Information System Department, School of Bioscience, Technology, and Innovation, Atma Jaya Catholic University of Indonesia, 15345, Jakarta, Indonesia.

[4]Management Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Jakarta, 11480, Indonesia.

*Corresponding Author's Email: christine.dewi@uksw.edu

Article History: Received 5 January 2026; Revised 10 April 2026; Accepted 19 April 2026

ABSTRACT: The transition toward Industry 4.0 has intensified the demand for zero-defect manufacturing, making Automated Visual Inspection a critical component of smart factories. However, implementing deep learning in real-world production faces significant challenges, particularly regarding data scarcity and edge deployment constraints. This paper provides a comprehensive review of YOLO-based applications in smart manufacturing, categorizing recent implementations across surface defect detection, assembly verification, robotic vision, and predictive maintenance. Our analysis reveals a critical paradigm shift from model-centric optimization toward data-centric strategies. While advanced architectures improve detection precision, integrating few-shot learning techniques and lightweight models is essential to overcome industrial data limitations. Ultimately, this review establishes a foundational roadmap for achieving vision-driven, zero-defect production in resource-constrained factory environments.

KEYWORDS: *Automated Visual Inspection; Defect Detection; YOLO; Industry 4.0; Edge Computing.*

1.0 INTRODUCTION

The transition toward Industry 4.0 has intensified the demand for zero-defect manufacturing, compelling industries to adopt advanced digital technologies and Artificial Intelligence (AI)-driven solutions. Conventional manual quality inspection, which is inherently subjective, labor-intensive, and prone to human error, is increasingly inadequate in meeting the precision and consistency requirements of modern smart factories [1]. Consequently, deep-learning-based Automated Visual Inspection (AVI) systems have become a core component of intelligent production lines, enabling real-time defect detection and continuous process monitoring [2].

Among various deep learning architectures, the You Only Look Once (YOLO) family has emerged as one of the most widely adopted object detection frameworks in industrial visual inspection. As a one-stage detector, YOLO offers a favorable trade-off between detection accuracy and real-time inference speed, making it highly suitable for integration with manufacturing equipment and robotic systems [3]. The architectural evolution of YOLO, from early anchor-based designs to anchor-free mechanisms and Transformer-integrated variants such as YOLOv11 and YOLO NAS, reflects ongoing efforts to enhance global feature representation and improve detection performance on small and complex industrial defects [3, 4]. YOLO-based approaches have been extensively applied to automotive inspection, precision component verification, and surface defect detection, including scratches, dents, and microcracks [2, 4, 5]. Its real-time capabilities position it as a critical enabler for intelligent quality control within cyber-physical production systems.

Despite promising performance reported in controlled experimental settings, a significant gap persists between laboratory-level achievements and industrial-scale deployment. Recent studies indicate an average adaptation delay of approximately three years between the introduction of state-of-the-art (SOTA) detection models and their adoption in manufacturing environments [1]. This delay is primarily attributed to infrastructure constraints, including the high computational demands of advanced YOLO variants and their limited compatibility with resource-constrained edge devices. Additionally, real-world industrial environments introduce complex visual noise and domain variability that often degrade model robustness and generalization performance [2, 4]. Despite the growing number of studies investigating YOLO-based visual inspection, existing research largely concentrates on improving detection accuracy or proposing architecture-specific enhancements without systematically examining their readiness for real industrial deployment. Current surveys typically address automated visual inspection or YOLO architectural evolution independently, while limited attention has been given to integrating defect detection, assembly verification, robotic vision, and predictive maintenance within a unified Industry 4.0 manufacturing perspective. Moreover, challenges related to edge deployment constraints, data scarcity in industrial environments, and scalability toward zero-defect production remain insufficiently synthesized across different application domains. To address these challenges, this review aims to bridge the gap between theoretical YOLO advancements and practical industrial constraints by synthesizing recent developments into a structured implementation roadmap for smart manufacturing.

The structure of this paper is as follows: Section 2 describes Related Work, which discusses the evolution of the YOLO architecture in industrial vision, the application taxonomy in smart manufacturing (surface detection, assembly verification, and robotic vision), and the technical challenges related to deployment and learning. Section 3 explains the systematic literature review methodology used, including search strategies and selection criteria. Section 4 presents Analysis and Results regarding model trends, application distribution, and in-depth analysis of edge deployment and learning paradigms. Section 5 presents a discussion and synthesis of cross-domain findings, and Section 6 concludes the paper with future work directions and implications for researchers and practitioners.

2.0 Evolution of YOLO in Industrial Vision

The progression of the You Only Look Once (YOLO) algorithm in industrial settings reflects a shift from generic object detection to specialized architectures optimizing the trade-off between speed, accuracy, and training stability. This evolution can be categorized into three distinct phases: the early anchor-based era, the optimization for efficiency, and the recent integration of advanced feature retention mechanisms.

As summarized by Terven et al. [3], early YOLO versions (YOLOv1 to YOLOv3) were developed under the Darknet framework, followed by later PyTorch-based implementations spanning from YOLOv5 to YOLOv8, and eventually evolving into architectures like YOLO-NAS. A key architectural transition occurred from anchor-based detectors (YOLOv2 to YOLOv7) to anchor-free mechanisms consolidated in YOLOv8. This transition aimed to reduce reliance on manually designed anchor boxes and improve adaptability to objects with varying aspect ratios. In industrial inspection contexts, Chen and Shiu [6] compared multiple YOLO generations for metallization quality assessment, reporting improved training stability in YOLOv5 under limited dataset conditions compared to earlier architectures.

To accommodate real-time requirements in manufacturing environments, several studies introduced lightweight modifications to existing YOLO backbones. For instance, Nikam et al. [7] evaluated multiple YOLO iterations to identify the most optimal model capable of detecting multi-scale defects in complex additive manufacturing processes without sacrificing inference speed. Building on this need for specialized detection, Shen et al. [8] adapted YOLOv5 for Printed Circuit Board Assembly (PCBA) inspection by incorporating larger receptive field components and alternative loss functions to address component-specific geometries. In parallel, Mao et al. [9] integrated YOLOv7 with Generative Adversarial Networks (ConSinGAN) to augment limited defect datasets, demonstrating improved detection performance under data-scarce conditions.

Recent developments emphasize improved feature retention for detecting fine-grained industrial defects. Further expanding on their previous analysis, Nikam et al. [7] explored YOLOv9 variants incorporating enhanced gradient information and layer aggregation mechanisms to preserve microscopic defect features during down sampling. Similarly, Yang et al. [4] proposed TMSS-YOLO based on YOLOv11, integrating transformer-based modules and spatial attention blocks to refine feature representation. Melgar et al. [10] further examined YOLO-World, which introduces zero-shot detection capabilities for anomaly identification in adaptive monitoring systems. To provide a clearer overview of this progression, Table 1 summarizes the key architectural upgrades of the most widely used YOLO variants in the manufacturing sector.

Table 1: Floating-point operations necessary to classify a sample

YOLO Variant	Key Architectural Upgrade	Target Industrial Application	Key Advantage in Manufacturing	Ref.
YOLOv5	Focus on training stability; lightweight backbone modifications.	Metallization quality, Printed Circuit Board Assembly (PCBA) inspection.	High stability under limited dataset conditions; efficient edge deployment.	[6], [8]
YOLOv7	Trainable "bag-of-freebies"; optimized for integration with GANs (e.g., ConSinGAN).	Mass-produced electronic components.	Excellent performance augmentation under data-scarce conditions.	[9]
YOLOv8	Transition to anchor-free mechanism.	General industrial defect detection.	Improved adaptability to defects with varying aspect ratios.	[3]
YOLOv9	Enhanced gradient information and layer aggregation.	Laser-directed energy deposition (Additive Manufacturing).	Superior preservation of microscopic/multi-scale defect features.	[7]
YOLOv11	Integration of Transformer modules and spatial attention blocks.	Industrial cutting tool detection.	Refined global feature representation and robust domain generalizability.	[4]
YOLO-World	Vision-language pre-training; zero-shot detection capabilities.	Adaptive anomaly identification in heavy machinery.	Eliminates the need for extensive retraining for new, unseen defects.	[10]

Based on the comparison in Table 1, it is evident that the architectural trajectory of YOLO models is increasingly aligning with the complex demands of smart manufacturing. The shift from anchor-based designs (such as YOLOv5 and YOLOv7) to anchor-free mechanisms and Transformer-integrated architectures (such as YOLOv8 and YOLOv11) demonstrates a significant improvement in handling multi-scale anomalies without the need for exhaustive manual parameter tuning. This evolution is particularly crucial for industrial visual inspection, as it enhances the model's capability to detect fine-grained and microscopic defects on complex surfaces like PCBAs or reflective metallization. Furthermore, the recent introduction of vision-language pre-training in YOLO-World marks a critical milestone for data-scarce environments, enabling zero-shot anomaly detection without requiring extensive dataset collection and retraining for new, unseen defects.

Ultimately, these structural optimizations provide a more favorable balance between high-precision feature extraction, which is the foundation of zero-defect production, and computational efficiency, thereby lowering the barrier for real-time edge deployment on the factory floor.

2.2 YOLO Application in Smart Manufacturing

2.2.1. Surface and structural defect detection

The application of YOLO algorithms in surface and structural defect inspection has expanded from laboratory experiments to domain-specific solutions. Recent literature on this topic generally focuses on handling material heterogeneity, selecting appropriate model versions, addressing class imbalance, and modifying architectures for specific industrial constraints.

To provide a clear overview of how these algorithms are deployed in real-world scenarios, Figure 1 illustrates the conceptual pipeline of a typical YOLO-based automated visual inspection system within a manufacturing line.

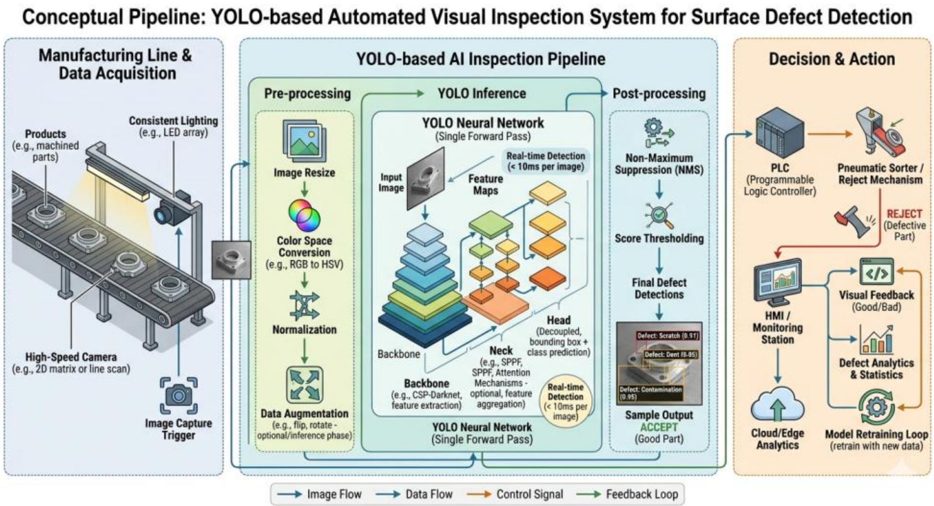

Figure 1: Conceptual pipeline of a YOLO-based automated visual inspection system for surface defect detection in a manufacturing line.

(a) Material characteristics and domain challenges

Distinct material properties necessitate tailored computer vision strategies. For reflective metallic materials, Guo et al. [11] and Xu [12] addressed dataset noise caused by irregular light reflections, while Yu et al. [13] focused on transparent polymer films where defects are visually subtle. In additive manufacturing and PCB assembly, challenges shift towards topological complexity. Nikam et al. [7] and Shen et al. [8] utilized datasets requiring the distinction of fine features, such as voids versus shadows, within dense backgrounds. Most studies rely on proprietary datasets from specific production lines, suggesting that cross-domain generalization across different material batches remains a relevant consideration [12].

(b) Model selection and adaptation

Researchers select YOLO versions based on the specific trade-off between stability and feature handling. Lin et al. [14] and Xu [12] retained the YOLOv4 architecture for its established stability. Conversely, recent studies [15, 16] have adopted YOLOv8, reporting that its anchor-free mechanism effectively handles defects with extreme aspect ratios. Furthermore, Nikam et al. [7] evaluated YOLOv9-GELAN, noting improved feature retention for complex additive manufacturing defects compared to earlier architectures.

(c) Strategies for class imbalance and small defects

To address the scarcity of defect samples, several approaches have been proposed. Yu et al. [13] implemented a semi-supervised framework, utilizing unlabeled data to enhance feature representation for micro defects. Alternatively, Yuan et al. [17] applied Inter-class Similarity Distillation (ISD) at the feature level to improve the detection of small defects on PCBs obscured by background noise. While these methods demonstrate metric improvements, fewer studies explicitly evaluate performance under open-world or zero-shot scenarios where novel defect types may emerge.

(d) Architectural modifications: efficiency vs. complexity

Modifications to the YOLO architecture typically follow two distinct directions. One approach prioritizes efficiency through parameter reduction; Xu [12], Yuan et al. [17], and Liu et al. [15] integrated lighter structures like GhostNet to reduce model size. The alternative approach focuses on enhancing feature extraction capability; Guo et al. [11] and Shen et al. [8] incorporated Transformer modules or large kernel convolution blocks to capture global context, a modification that increases computational load but aims to improve detection on large surfaces.

(e) Performance under industrial conditions
While architectural modifications frequently result in high reported mAP and FPS in controlled experiments [8], real-world deployment faces variables such as machine vibrations and lighting fluctuations. Successfully isolating surface defects is often a precursor to broader integration challenges involving hardware constraints and data scarcity, which are discussed in subsequent sections.

2.2.2. Data-efficient strategies in assembly and robotic vision

A persistent challenge in manufacturing deep learning is the scarcity of high-quality, annotated defect data despite the abundance of operational data. Hütten et al. [1] and Ma et al. [2] note that class imbalance and annotation costs severely hinder industrial deployment, particularly for automated assembly lines and robotic vision systems that require robust generalization. Consequently, recent literature has explored various data engineering strategies and learning paradigms to mitigate these constraints.

To leverage the volume of unlabelled product images available on factory floors, semi-supervised frameworks have also been adopted. Yu et al. [13] implemented a teacher-student mechanism within a YOLOv8 architecture for detecting micro-defects in industrial polymer films. By generating pseudo-labels from unlabelled data, this method attempts to expand the learned data distribution beyond the limitations of manually annotated sets, particularly for defects that are difficult to label consistently.

Furthermore, robotic vision systems in agile production environments often require rapid adaptation to novel defect types or complex manipulation tasks. Zhang et al. [18] addressed this by proposing the Small Object Few-Shot (SOFS) model, which employs precise feature weighting to segment anomalies using a minimal number of support images, overcoming the loss of spatial information often associated with resizing in conventional segmentation.

Distinguished from model-centric optimization, some research focuses on enhancing input data quality through data-centric engineering. Im et al. [19] investigated the impact of preprocessing techniques, such as Background Removal (BR) and Contrast Limited Adaptive Histogram Equalization (CLAHE), within the plastic injection molding industry. Their findings suggest that rigorous data cleaning and histogram balancing can improve detection robustness against environmental variations, offering a potential alternative to increasing architectural complexity.

2.3 Industrial Deployment and Learning Challenges

2.3.1. Edge deployment and hardware constraints

The deployment of YOLO models in manufacturing environments introduces additional constraints beyond those encountered in controlled laboratory settings. Standard YOLO architectures require considerable memory and processing resources, which are not always available on low-cost industrial cameras or embedded systems [20]. Studies have examined the latency characteristics of full-scale models when deployed on edge devices, particularly in high-speed production line contexts [17, 21].

Several studies explore model compression and hardware-specific optimization to enable edge deployment. Lightweight architectures, such as Tiny-YOLO and Nano variants, have been investigated as approaches to reducing inference time while maintaining detection capability [22]. Furthermore, edge deployment heavily emphasizes real-time capabilities. In high-speed production lines, the system's throughput, measured in frames per second (FPS), must account for the entire pipeline latency rather than just the model execution. The actual FPS achieved on the factory floor is calculated as in Eq. (1):

$$FPS = \frac{1}{(t_{preprocess} + t_{inference} + t_{postprocess})}) \quad (1)$$

where t_preprocess, t_inference, and t_postprocess represent the time taken (in seconds) for image preparation, neural network execution, and bounding box filtering (e.g., Non-Maximum Suppression), respectively. Minimizing this total latency is crucial for maintaining synchronization with manufacturing equipment.

However, while optimizing for high FPS and reducing model size, it is imperative to ensure that the model's detection accuracy is not severely compromised. In industrial anomaly detection, avoiding false negatives (missed defects) and false positives (false alarms) is evaluated using Precision (P) and Recall (R), as formulated in Eq. (2):

$$P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN} \quad (2)$$

where TP, FP, and FN denote True Positives, False Positives, and False Negatives. The standard metric evaluated alongside FPS to measure overall object detection performance is the mean Average Precision (mAP), calculated as in Eq. (3):

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (3)$$

where N is the total number of defect classes, and AP_i represents the Average Precision (the area under the Precision-Recall curve) for the i-th class. The primary challenge in edge deployment is finding the optimal balance between computational efficiency (high FPS) and detection reliability (high mAP). The deployment of YOLO on Field-Programmable Gate Arrays (FPGAs) has also been examined as a method for achieving low-latency inference under power-constrained conditions [23]. A study on wood defect detection demonstrated that edge-deployable models can operate under limited computational resources while successfully maintaining high detection rates [24].

2.3.2. Data scarcity and class imbalance

The availability of high-quality labelled training data represents a recurring challenge in industrial computer vision applications. In production environments with low defect rates, defective samples are naturally rare, resulting in datasets where non-defective (normal) samples substantially outnumber defective ones [1, 15]. This distributional imbalance has been associated with reduced detection sensitivity toward minority defect classes in standard YOLO training configurations, often leading to model bias and increased false negative rates [25].

Data augmentation and synthetic data generation have been extensively examined as primary approaches to increasing the representation of defective samples in training datasets [9, 26]. Traditional augmentation techniques (e.g., rotation, scaling, and color jittering) are often insufficient to capture the complex topological variations of real-world defects. Consequently, Generative Adversarial Networks (GANs) have been applied to produce highly realistic synthetic defect images, providing a crucial supplementary source of training samples in contexts where physical defect collection is severely constrained [9].

The core mechanism of GAN-based defect synthesis involves a zero-sum game between two neural networks: a Generator (*G*) that synthesizes fake defect images from random noise, and a Discriminator (*D*) that evaluates whether an image is real (from the factory dataset) or fake. To provide a clearer visualization of this adversarial training process, Figure 2 illustrates the conceptual architecture of a standard GAN.

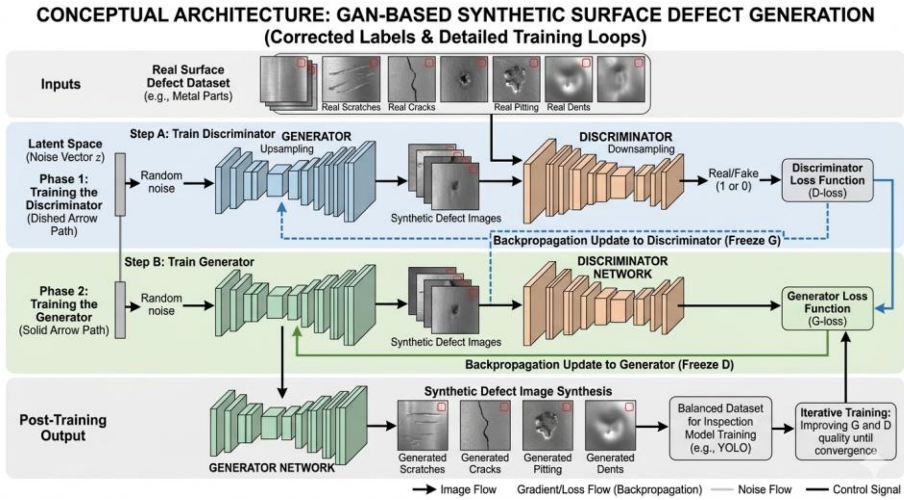

Figure 2: Conceptual architecture of a Generative Adversarial Network (GAN) for synthetic defect image generation.

To enrich the minority class distribution, the models are trained simultaneously using the standard minimax objective function, as formulated in Eq. (4):

$$V(D, G) = E_{x\sim Pdata(x)}[log\ \ D(x)] + E_{z\sim P_z(z)}[log\ \left(1 - D\left(G(z)\right)\right)] \quad (4)$$

where x represents the real defective samples from the data distribution P_{data}, and z represents the random noise input from the prior distribution P_z . By optimizing Equation (4), studies like Mao et al. [9] successfully forced the Generator to produce high-fidelity defect samples (e.g., missing pins or scratches) that are statistically indistinguishable from genuine anomalies. This synthetic augmentation allows YOLO models to learn robust boundary features for minority classes without requiring months of defective data collection on the factory floor.

2.3.3. Predictive maintenance and process monitoring

YOLO-based vision systems have been applied beyond product inspection toward monitoring the operational condition of manufacturing equipment. These applications focus on detecting changes in machinery state, including tool wear progression and structural anomalies during active operation [4, 10]. Integrating object detection into maintenance workflows allows for a transition from reactive to proactive strategies, minimizing unplanned downtime in smart factories.

Recent studies have highlighted the versatility of YOLO in identifying diverse failure modes in heavy machinery and precision tools. For instance, Melgar et al. [10] demonstrated that YOLO-variants, including YOLOv11 and YOLO-World, can effectively monitor critical components by detecting hose wear, piston failures, corrosion, and moisture under variable industrial lighting conditions. Similarly, Yang et al. [4] utilized an enhanced YOLOv11 framework, termed TMSS-YOLO, to classify the wear state of industrial cutting tools. Their study emphasizes that high-precision detection of tool degradation is vital for maintaining machining accuracy and preventing catastrophic equipment failure.

Furthermore, studies involving laser-directed energy deposition (L-DED) processes have demonstrated the use of YOLO for real-time process stability monitoring. Nikam et al. [7] evaluated YOLOv7, YOLOv8, and YOLOv9 to detect flash formation, voids, and rough textures, identifying deviations associated with equipment condition changes during the additive manufacturing process. Unlike static surface inspection, these approaches require the model to handle complex industrial backgrounds to ensure synchronization with production requirements. By extending YOLO applications toward equipment condition monitoring, manufacturers can achieve a more holistic integration of vision systems within cyber-physical production environments, ultimately supporting the goals of autonomous zero-defect manufacturing.

3.0 METHODOLOGY

This study conducts a structured thematic literature review focusing on the implementation of You Only Look Once (YOLO) algorithms within the context of Industry 4.0 smart manufacturing environments. The primary objective is to systematically analyze recent advancements in YOLO architectures, including YOLOv5, YOLOv7, YOLOv8, YOLOv10, YOLOv11, and YOLOv12 variants, and their application across four major industrial domains: defect detection, assembly verification, robotic vision, and predictive maintenance. The growing adoption of deep learning-based automated visual inspection systems in manufacturing [1], together with the increasing research focus on industrial surface defect detection [2], motivates the need for a structured and thematic synthesis. Unlike conventional narrative reviews, this study adopts a semi-systematic approach to ensure methodological rigor while maintaining flexibility in thematic classification.

The literature search was conducted across major scientific databases, including IEEE Xplore, ScienceDirect, and Google Scholar. To ensure the inclusion of contemporary technological developments, the search was restricted to publications from 2021 to 2025, a period marked by rapid architectural evolution in YOLO-based systems and their industrial deployment [3, 4, 11]. The search strategy employed a structured combination of keywords, specifically ("YOLO" OR "You Only Look Once") AND ("Industry 4.0" OR "Smart Manufacturing" OR "Defect Detection" OR "Visual Inspection" OR "Robotic Vision" OR "Predictive Maintenance"). This formulation was designed to capture studies explicitly linking YOLO-based detection with industrial automation contexts, as demonstrated in defect detection for electronics manufacturing [8, 9], industrial aluminium sheets [12], weld inspection [15], and robotic applications [28]. To systematically select the most relevant research, this study employed a comprehensive workflow encompassing literature search, identification, and a rigorous filtering process.

Following the initial identification of records, a multi-stage screening procedure was applied to ensure a high-quality synthesis. Articles were excluded if they focused on non-industrial domains such as agriculture, medical imaging, or autonomous driving without direct manufacturing relevance. Studies were also excluded if they lacked quantitative experimental validation, since industrial inspection research consistently emphasizes measurable metrics such as mean

Average Precision, inference speed, and hardware deployment feasibility [20, 23, 24]. Furthermore, purely theoretical architectural proposals without implementation on industrial datasets were not considered, in alignment with the practical orientation of sustainable smart manufacturing research [25].

To ensure conceptual clarity, the scope of this review is explicitly defined. This study focuses exclusively on YOLO-based object detection frameworks applied within industrial production environments. It does not aim to provide a comprehensive comparison between YOLO and other detection architectures beyond contextual reference, as the focus is on understanding YOLO's role within smart manufacturing ecosystems. Additionally, this review does not deeply explore semantic segmentation-based inspection systems unless they are directly integrated with YOLO detection pipelines. By narrowing the scope to industrial smart manufacturing, this study ensures thematic coherence and analytical depth consistent with contemporary automated visual inspection research [1, 2].

After the screening and eligibility assessment, a total of 30 highly relevant studies were selected as the core references of this review. Each selected study was systematically examined using a structured data extraction strategy. The extraction process was guided by predefined analytical dimensions to enable cross-study comparison and thematic synthesis. For each paper, detailed information was recorded regarding the YOLO version utilized, architectural modifications such as attention mechanisms [23], transformer integration [11], lightweight backbone replacement [17], and edge-oriented optimization strategies [20], deployment environment including GPU-based training, FPGA acceleration [23], or edge-device implementation [24], dataset characteristics such as defect type and scale variability [13], and reported performance metrics including mAP, precision, recall, inference speed, and model complexity.

Beyond quantitative metrics, qualitative attributes were also considered during data extraction. These include the degree of real-world industrial validation [25], adaptability to small-object detection scenarios such as micron-scale defects [13], and scalability to production-scale environments. The inclusion of few-shot learning approaches for small-object industrial inspection [18] further informed the analytical framework, particularly in addressing data scarcity challenges common in manufacturing contexts.

To enable structured synthesis, the selected studies were categorized into four thematic domains: defect detection, assembly verification, robotic vision, and predictive maintenance. This classification reflects dominant application areas identified in automated visual inspection literature [1, 2], as well as domain-specific industrial implementations such as robot grasping [28] and predictive maintenance systems [10]. A secondary classification captured cross-cutting architectural strategies including lightweight optimization [17], edge deployment feasibility [20], semi-supervised learning [13], and attention-enhanced YOLO models [23]. This layered classification framework supports analytical interpretation of technological convergence trends rather than isolated performance comparisons.

The semi-systematic approach adopted in this study balances structured methodological rigor with analytical flexibility. Rather than merely aggregating reported performance values, this review emphasizes architectural evolution [3], industrial deployment feasibility [24], data scarcity challenges [18], and the integration of YOLO-based visual intelligence within sustainable smart manufacturing systems [25].

4.0 RESULT

4.1 Publication Trends

To understand the trajectory of object detection in Industry 4.0, we analysed the temporal distribution of the 30 selected papers published from 2021 to 2025. Fig. 3 illustrates the annual publication volume of these reviewed articles.

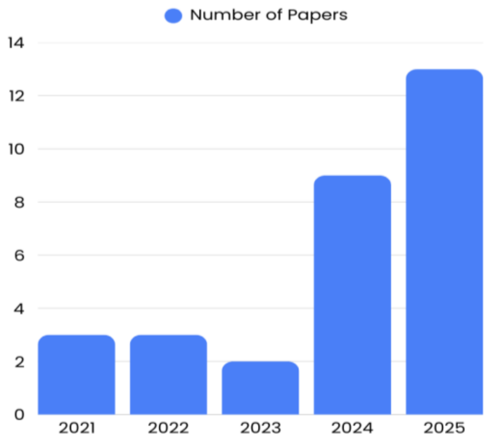

Figure 3: Annual distribution of selected publications (2021–2025).

As shown in Fig. 3, research interest remained relatively stable during the earlier period between 2021 and 2023. During this time, early iterations of YOLO, such as YOLOv4 and YOLOv5, were being adapted for fundamental industrial inspection tasks.

However, a prominent spike occurred starting in 2024, continuing with an even sharper surge into 2025. The year 2024 accounted for nine studies, while 2025 reached a peak of 13 publications, representing approximately 43.33% of the total selected studies. This rapid growth directly correlates with the release of user-friendly, high-performance models like YOLOv8 and the structural shift toward anchor-free mechanisms, which significantly improved the detection of microscopic or irregularly shaped defects.

Furthermore, this software evolution was supported by the increasing maturity of edge AI hardware (e.g., NVIDIA Jetson Orin, Raspberry Pi 5), allowing these deep learning models to be seamlessly integrated into real-time production lines. The overwhelming concentration of research in 2024 and 2025 clearly reflects the industry's urgent and accelerating demand for automated, zero-defect manufacturing solutions.

4.2 Comparative performance analysis

A comprehensive comparison of representative studies is presented in

Table 2. This summary highlights the domain-specific performance metrics, demonstrating how different YOLO variants adapt to varying industrial constraints.

Table 2: Performance comparison of YOLO variants on various edge computing hardware.

Application / Domain	Method / YOLO Variant	Dataset	Key Results & Strategy	Performance Metric (%)	Ref.
Automated Visual Inspection (Review)	YOLO (Various)	N/A (Review)	Comprehensive survey on deep learning for manufacturing and maintenance.	N/A (Focus on literature review)	**[1]**
Surface Defect Inspection (Review)	YOLO (Various)	N/A (Review)	Systematic review highlighting challenges in multi-scale and small object detection.	N/A (Focus on literature review)	**[2]**
Computer Vision Architecture (Review)	YOLOv1 to YOLOv8	N/A (Review)	Analyzed the evolution of YOLO architectures and the shift toward anchor-free mechanisms.	N/A (Focus on architectural evolution)	**[3]**
Industrial Cutting Tools	TMSS-YOLO (YOLOv11)	Industrial Tool Dataset	Redesigned Transformer block for global feature extraction and multi-scale fusion.	mAP: 94.98%, Accuracy: 95.71%	**[4]**
Surface Defect Detection	YOLOv1 to YOLOv8	Surface Defect Dataset	Optimized detection in complex industrial environments.	N/A (Focus on environmental robustness)	**[5]**
ABS Metallization	YOLOv2, v3, v4, v5	508 Electroplating Images	YOLOv5 offered the best balance for reflective product surfaces.	Accuracy: >70.0%	**[6]**
Laser-Directed Additive Manufacturin g	YOLOv9	Additive Manufacturing Images	Superior preservation of microscopic/multi-scale defect features using enhanced gradients.	N/A (Focus on microscopic defects)	**[7]**
PCB Assembly Inspection	YOLOv5	PCB Defect Dataset	High stability under limited dataset conditions for quality control.	N/A (Focus on edge efficiency)	**[8]**
Electronics (DIP Switch)	YOLOv7 + ConSinGAN	Real + GAN Generated	GAN augmentation outperformed traditional thresholding methods in data-scarce conditions.	Accuracy: 95.5%	**[9]**
Heavy Machinery Components	YOLOv11, RT-DETR, YOLO-	Industrial Images	YOLOv11 achieved the highest accuracy for early fault detection (predictive	mAP@0.5: 83.8%	**[10]**

	World		maintenance).		
Steel Surface Defect	MSFT-YOLO	NEU-DET Dataset	Integrated Transformer blocks to focus on subtle texture defects on metal surfaces.	N/A (Focus on feature extraction)	**[11]**
Industrial Aluminum Sheets	YOLOv4 + GhostNet	Aluminum Sheet Defects	Lightweight model replacing CSPDarknet53 with GhostNet to reduce parameter size.	mAP: 98.1%	**[12]**
Industrial Polymer Films	Enhanced YOLOv8	Polymer Film Images	Semi-supervised framework implemented for micron-scale defect detection.	N/A (Focus on semi-supervised learning)	**[13]**
Pharmaceutical Production	YOLOv8	Lyophilized Powder Data	Vision-based inspection system for detecting foreign objects in production lines.	N/A (Focus on pharmaceutical safety)	**[14]**
Automotive (Spot Weld)	TA-YOLOv8	Real Production Data	Lightweight model reduced parameters by 32% for efficient edge systems deployment.	N/A (Focus on relative improvement)	**[15]**
Textile / Fabric Defect	YOLOv5	Fabric Image Dataset	Improved detection for woven fabric defects on continuous production lines.	N/A (Focus on real-time textile inspection)	**[16]**
Printed Circuit Boards (PCB)	LW-YOLO	PCB Defect Dataset	Lightweight deep learning model designed for fast and precise PCB defect detection.	mAP: 96.4%	**[17]**
Vision-Based Industrial Inspection	SOFS (YOLO-based)	Few-Shot Industrial Dataset	Small Object Few-shot Segmentation to locate unseen defects with minimal annotations.	mIoU: 65.2%	**[18]**
Plastic Injection Molding	Data-Centric YOLOv4	Self-created (1,534 Images)	Resolved false positives from light reflection using background removal and CLAHE.	N/A (Focus on data preprocessing)	**[19]**
Visual Conformity	YOLOv8n	Conformity Dataset	Natively lightweight model deployed directly on low-cost edge devices (Raspberry Pi 4).	N/A (Focus on hardware deployment)	**[20]**
Industrial Surface Inspection	Edge-YOLO	Surface Anomaly Dataset	Edge-enhanced backbone to retain shallow network features for minute defect detection.	N/A (Focus on edge deployment)	**[21]**

Manufacturing Systems	Tiny YOLOv4	General Manufacturing	Learning computer vision systems using tiny architectures for industrial environments.	N/A (Focus on tiny architectures)	**[22]**
Auto Parts Defect (FPGA)	YOLOv3 + Attention	40,320 Images	Hardware acceleration on FPGA (Zynq-7020) with ultra-low power consumption.	Accuracy: 99.2%	**[23]**
Wood Defect Detection	YOLOv5/v8	Wood Material Dataset	Edge-deployable model for sustainable and smart manufacturing of wood products.	N/A (Focus on sustainability)	**[24]**
Sustainable Smart Manufacturing	Deep Learning (YOLO)	Manufacturing Dataset	Deep learning framework designed for sustainable smart manufacturing environments.	N/A (Focus on sustainable frameworks)	**[25]**
Integrated Circuit Board	YOLO + Image Augmentation	IC Board Dataset	Image augmentation fusion model to improve detection robustness on circuit boards.	N/A (Focus on data augmentation)	**[26]**
Wind Turbine Blades	YOLO-WTB (YOLOv12n)	Aerial Imagery	Detected micro-cracks on small objects using enhanced, ultra-modern architecture.	N/A (Focus on architectural capability)	**[27]**
Robotic Grasping (Industry 4.0)	YOLOv4, YOLOv7	RGB & RGB-D Grasping	The system successfully grasped objects under occlusion, prioritizing the most visible objects.	N/A (Focus on robotic vision)	**[28]**
Occupational Safety (PPE)	YOLO	Teaching Lab Dataset	Detected proper utilization and adequacy of Personal Protective Equipment (PPE).	mAP: 75.7%	**[29]**
Autonomous Logistics / AGV	Tiny-YOLOv4	Campus Road (2,000 Images)	Embedded system for obstacle avoidance and communication in autonomous vehicles.	Accuracy: 83.0%	**[30]**

Based on the comparison in Table 2, it is evident that YOLO architectures have been successfully adapted across a wide spectrum of Industry 4.0 applications, ranging from high-precision electronic manufacturing [17, 26] to heavy industrial inspections [16, 21, 27] and robotic vision [28, 30]. A clear trend emerges regarding the trade-off between model complexity and deployment constraints. For instance, studies targeting edge devices and embedded systems often favor natively lightweight networks; research on visual conformity utilized YOLOv8n for low-cost monitoring on Raspberry Pi 4 [20], while Tiny-YOLOv4 was implemented for embedded obstacle avoidance and communication in autonomous operations [30].

In contrast, for hardware-accelerated deployment, such as Field-Programmable Gate Arrays (FPGA), structural optimization of established models proves highly effective. For example, a study focusing on auto parts defect detection demonstrated this by integrating an attention mechanism into a YOLOv3 architecture [23]. Through hardware acceleration on an FPGA (Zynq-7020) platform, this implementation achieved an impressive 99.2% accuracy with ultra-low power consumption, making it highly suitable for resource-constrained manufacturing environments.

Furthermore, domains requiring the detection of microscopic or subtle defects, such as Printed Circuit Board (PCB) inspection and wind turbine blades, increasingly integrate advanced architectures like LW-YOLO [17] and YOLO-WTB (based on YOLOv12n) [27], with the former reaching a mean Average Precision (mAP) of 96.4%. Finally, the table highlights a significant shift towards data-efficient paradigms. Rather than solely optimizing the network depth, recent research emphasizes solving industrial data scarcity through image augmentation fusion [26], Small Object Few-shot Segmentation (SOFS) [18], and data-centric preprocessing techniques to resolve environmental noise like light reflection [19]. This collective evidence suggests that the future of industrial object detection lies not merely in deploying deeper networks, but in developing hardware-aware, data-efficient models tailored to specific manufacturing environments.

4.3 Application-specific findings

The review identifies two major clusters of YOLO applications in smart manufacturing: surface defect detection and real-time edge monitoring.

4.3.1. Edge surface defect detection

Surface inspection [31, 32] remains the most dominant application of YOLO in the manufacturing sector [33, 34]. The review highlights that the primary challenge in this domain is the detection of small objects against complex backgrounds, such as scratches on metal or misaligned components on printed circuit boards (PCBs) [35, 36, 37].

Findings from the reviewed papers indicate that standard YOLO models often struggle with these small defects due to the loss of feature information in deep convolutional layers. To address this, researchers have successfully integrated attention mechanisms and feature pyramid networks (FPNs). For instance, studies on PCB inspection [17] utilized lightweight modifications to enhance feature extraction without significantly increasing the computational load. Furthermore, data scarcity for rare defects is a recurring issue. Several papers address this by employing generative adversarial networks (GANs) or synthetic data generation [9] to augment training sets, thereby enabling few-shot learning capabilities where the model learns from limited defect samples.

4.3.2. Real-time monitoring on edge devices

The second critical finding relates to the deployment of YOLO on edge devices such as NVIDIA Jetson Nano, Raspberry Pi, and field-programmable gate arrays (FPGAs) [20, 23]. In Industry 4.0, cloud computing often introduces latency that is unacceptable for high-speed production lines.

The review reveals that model pruning and quantization are the most effective strategies for edge deployment. Papers focusing on tiny versions of YOLO (e.g., Tiny-YOLOv4, YOLOv5n) demonstrate that it is possible to achieve real-time inference speeds (greater than 30 frames per second) on embedded systems, albeit with a slight trade-off in accuracy. For example, in autonomous logistics applications [30], the use of Tiny-YOLO enabled successful obstacle avoidance with minimal processing delay. Additionally, recent implementations on Raspberry Pi [20] confirm that modern lightweight YOLO variants can operate effectively on low-cost hardware, supporting the trend toward green AI in manufacturing. To provide a comprehensive overview of how different YOLO architectures perform under various hardware constraints, Table 3 summarizes the inference speeds and optimization strategies across selected edge deployment studies.

Table 3: Performance comparison of YOLO variants on various edge computing hardware.

Application Domain	YOLO Variant	Optimization Strategy	Edge Hardware	Inference Speed (FPS)	Ref.
Auto Parts Defect	YOLOv3 (Improved)	Pruning (14 layers removed)	FPGA (Zynq-7020)	High (Low latency, 10 W)	[23]
Visual Conformity	YOLOv8n	Natively lightweight	Raspberry Pi 4	~10–15 FPS	[20]
Autonomous Logistics	Tiny-YOLOv4	TensorRT acceleration	NVIDIA Jetson Nano	> 30 FPS	[30]
Spot Weld Anomaly	TA-YOLOv8	Parameter reduction (32%)	Edge Server / IPC	Real-time capable	[15]

As detailed in Table 3, the choice of edge hardware significantly dictates the required model optimization strategy. While FPGA implementations [23] excel in power efficiency through deep structural pruning, GPU-accelerated embedded devices like the NVIDIA Jetson Nano [30] effectively leverage TensorRT acceleration to maximize frame rates for dynamic autonomous tasks. Furthermore, the successful deployment of YOLOv8n on a standard Raspberry Pi 4 [20] demonstrates that natively lightweight models can provide a highly cost-effective solution for visual conformity tasks where extremely high-speed processing is not the primary requirement. Ultimately, this hardware-software co-design is essential for realizing scalable, green, and sustainable smart manufacturing systems

5.0 DISCUSSION

5.1 Architectural Evolution and Industrial Readiness

The development of the YOLO architecture in an industrial context shows a clear shift from the initial anchor-based approach to a more efficient and modular design. A comprehensive review by Terven et al. [3] confirms the evolution from YOLOv1 [38] to YOLOv8 [39] and YOLO-NAS [40], which focuses on improving training stability, inference efficiency, and the accuracy of small object detection [41, 42]. In the manufacturing domain [43, 44], implementational studies such as Chen and Shiu [6] show that newer generations of YOLO provide a better balance between accuracy and stability compared to previous versions.

In surface inspection applications, various architectural modifications have been developed to improve sensitivity to small defects and complex textures. The integration of transformers in MSFT-YOLO [11] and YOLOv11-based TMSS-YOLO [4] demonstrates improved global feature representation capabilities. Similarly, lightweight models such as TA-YOLOv8 [15] and LW-YOLO [17] are designed to reduce parameter complexity without sacrificing detection precision. The strategic divergence in these architectural developments is illustrated in Fig. 4, which classifies the models based on their primary optimization goals.

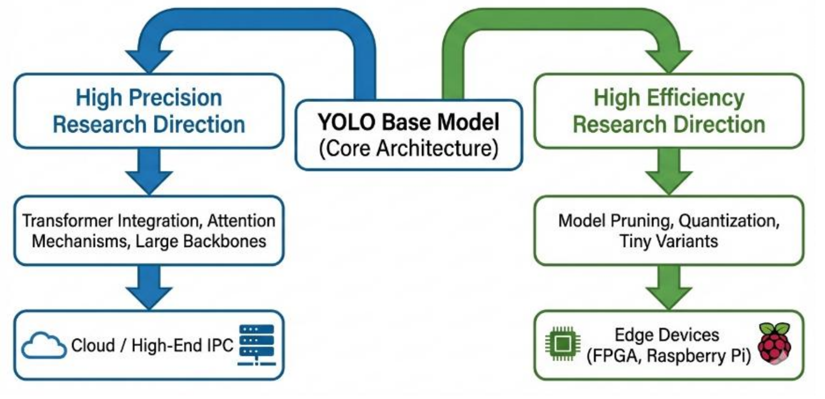

Figure 4: Divergent evolutionary paths of YOLO architectures in smart manufacturing: feature enhancement versus edge-oriented optimization.

As depicted in Fig. 4, the evolution of YOLO follows two distinct trajectories: one toward high-precision feature enhancement for complex defect identification, and another toward lean, edge-oriented optimization for real-time deployment. However, increased architectural complexity does not always align with industrial infrastructure readiness. Hütten et al. [1] and Ma et al. [2] emphasize the critical gap between laboratory performance and real-world implementation in production environments. A major challenge arises from the high computational requirements of the latest state-of-the-art models, which are often incompatible with low-cost edge devices. In response to these limitations, various studies propose edge-aware approaches.

Implementations of YOLO on Raspberry Pi [20, 45, 46, 47] and FPGA [53], alongside the development of Edge-YOLO [51, 48, 49] and CFIS-YOLO [24], show that model optimization through pruning, quantization, and parameter reduction are key strategies for maintaining real-time performance. Tiny-YOLOv4 [50] has also proven highly effective in embedded system scenarios involving autonomous tasks [22, 30, 51, 52]. Thus, as illustrated by current literature, while the evolution of the YOLO architecture has drastically improved detection accuracy, the main limiting factors in the context of Industry 4.0 are no longer model performance alone, but rather the readiness of edge infrastructure and computational efficiency.

5.2 Data Centric Challenges and Future Directions

In addition to architectural aspects, the most consistent challenge in industrial visual inspection is data limitations. Unbalanced class distribution and scarcity of defect samples are major obstacles in training detection models [1, 25]. A systematic study by Ma et al. [2] confirms that data scarcity is a structural issue in industrial defect inspection. Several studies have attempted to overcome these limitations through augmentation and semi-supervised learning approaches. Yu et al. [13] developed a YOLOv8-based semi-supervised framework for micro-defect detection, while Zhang et al. [18] introduced a few-shot segmentation approach for scenarios with limited samples. The integration of image augmentation on PCBs [26] also showed improved robustness against data variation.

Data-centric approaches are also an alternative to increasing architectural complexity. Im et al. [19] showed that preprocessing optimization and data quality engineering can improve inspection system stability without the need to enlarge the model. Meanwhile, applications in different domains such as additive manufacturing [7, 16, 61, 62], aluminium sheets [12, 63, 64, 65], and PCB assembly [8, 9, 66, 67] show that cross-domain generalization remains a significant challenge.

Extensions of YOLO applications towards predictive maintenance [10, 53], robotic grasping [28, 54, 55], and safety compliance [29, 56, 57] demonstrate the expanding role of visual detection in intelligent production systems [58, 59, 60]. However, most studies still focus on closed-set evaluation, with little exploration of open-world defect detection or adaptability to new defect types. Overall, the research direction shows a shift from model-centric optimization towards data-centric strategies that emphasize robustness, generalization, and training efficiency. Going forward, research needs to prioritize the integration of lightweight architecture and adaptive learning frameworks to ensure the sustainability of YOLO implementation in edge-based manufacturing systems.

Figure 5 provides a comprehensive multi-phase roadmap illustrating the evolution of YOLO-based visual inspection toward a vision-driven, zero-defect production paradigm. Phase 1, representing the current state, is meticulously optimized for high-performance, real-time edge deployment, utilizing lightweight architectures, model pruning, and quantization with robust high-throughput accelerators. Following this, Phase 2 delineates the emerging state, which shifts focus to addressing critical data scarcity and improving adaptability through advanced techniques like generative adversarial learning, diffusion models, and adversarial training. Collectively, this visual summary provides a clear research trajectory, demonstrating the transition from efficient deployment to increased model robustness and generalizability in complex manufacturing environments.

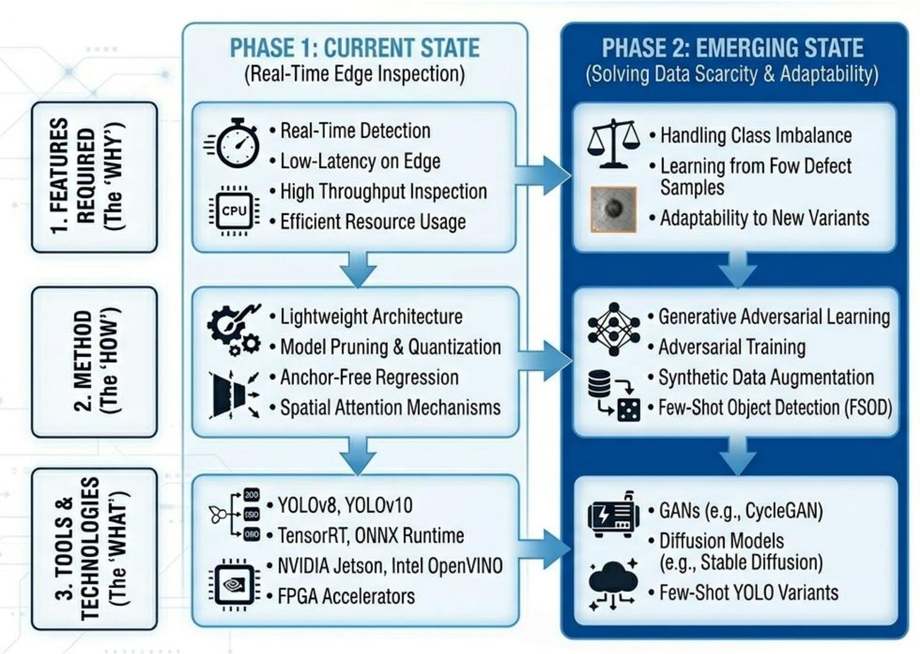

Figure 5: Evolution of YOLO based inspection.

6.0 CONCLUSION

This comprehensive review highlights the transformative impact of YOLO-based visual inspection systems in the context of Industry 4.0. By systematically analysing current literature, we categorized the applications of YOLO into four primary domains: surface defect detection, assembly verification, robotic vision, and predictive maintenance. While the continuous architectural evolution of YOLO algorithms has consistently delivered high detection accuracy and real-time processing capabilities, our findings indicate that architectural complexity is no longer the sole bottleneck. Instead, we identified a necessary paradigm shift toward data-centric strategies to address the chronic scarcity and imbalance of industrial defect samples.

Moving forward, the successful implementation of smart manufacturing will depend heavily on overcoming infrastructure limitations. Future research should prioritize the integration of synthetic data generation and few-shot learning frameworks to build robust models from limited data. Furthermore, developing lightweight architectures optimized for low-cost edge deployment will be crucial.

By focusing on these interconnected areas, data quality engineering and edge-aware model compression, industries can sustainably achieve vision-driven, zero-defect production in highly constrained factory environments.

AUTHOR CONTRIBUTIONS

C. Dewi and E. K. Pradibta Fury: Investigation, Data curation, Writing - review & editing, Validation, Writing – original draft, Formal analysis, Conceptualization. J. S. Rosario Putra and E. J. Feodora Aritonang: Methodology, Data curation, Writing – original draft, Writing – review & editing, Conceptualization. A. B. Setia Permana and M.H.F. Md Fauadi: Writing – original draft, Methodology, Data curation, Resources, Software, Investigation. D. Riantama: Visualization, Data curation, Writing – review & editing, Software, Writing –original draft, Investigation, Methodology. S. Aprius Sutresno: Visualization, Data curation, Writing – review& editing, Software, Writing – original draft, Investigation, Methodology.

ACKNOWLEDGMENT

The authors would like to express their sincere appreciation to Satya Wacana Christian University, Universiti Teknikal Malaysia Melaka, Universitas Ciputra, Atma Jaya Catholic University of Indonesia, and Bina Nusantara University for their academic and administrative assistance throughout the completion of this research.

REFERENCES

[1] N. Hütten, M. Alves Gomes, F. Hölken, K. Andricevic, R. Meyes, and T. Meisen, "Deep Learning for Automated Visual Inspection in Manufacturing and Maintenance: A Survey of Open-Access Papers," *Applied System Innovation*, vol. 7, no. 1, pp. 1-38, 2024.

[2] Y. Ma, J. Yin, F. Huang, and Q. Li, "Surface defect inspection of industrial products with object detection deep networks: a systematic review," *Artificial Intelligence Review*, vol. 57, pp. 1-48, 2024.

[3] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, pp. 1680–1716, 2023.

[4] G. Yang, Z. X. Zhang, H. D. Sang, H. Tang, and M. Wang, "TMSS-YOLO: An industrial tool detection method based on improved YOLOv11," *Signal, Image and Video Processing*, vol. 19, pp. 1-11, 2025.

[5] C. Gheorghe, M. Duguleana, R. G. Boboc, and C. C. Postelnicu, "Analyzing Real-Time Object Detection with YOLO Algorithm in Automotive Applications: A Review," *Computer Modeling in Engineering and Sciences*, vol. 141, pp. 1939–1981, 2024.

[6] Y. W. Chen and J. M. Shiu, "An implementation of YOLO-family algorithms in classifying the product quality for the acrylonitrile butadiene styrene metallization," *The International Journal of Advanced Manufacturing Technology*, vol. 119, pp. 8257–8269, 2022.

[7] D. Nikam, A. Chukwuemeke, A. Nigam, T. Bhosale, and S. Nikam, "On the application of YOLO-based object detection models to classify and detect defects in laser-directed energy deposition process," *Progress in Additive Manufacturing*, vol. 10, pp. 7609–7624, 2025.

[8] M. Shen, Y. Liu, J. Chen, K. Ye, H. Gao, J. Che, Q. Wang, H. He, J. Liu, Y. Wang, and Y. Jiang, "Defect detection of printed circuit board assembly based on YOLOv5," *Scientific Reports*, vol. 14, pp. 1–16, 2024.

[9] W. L. Mao, C. C. Wang, P. H. Chou, and Y. T. Liu, "Automated Defect Detection for Mass-Produced Electronic Components Based on YOLO Object Detection Models," *IEEE Sensors Journal*, vol. 24, pp. 26877–26888, 2024.

[10] P. C. Melgar, L. H. Chambi, and R. S. Torres, "Predictive Maintenance Based on Deep Learning: Early Identification of Failures in Heavy Machinery Components," *International Journal of Advanced Computer Science and Applications*, vol. 16, pp. 781–788, 2025.

[11] Z. Guo, C. Wang, G. Yang, G. Li, and H. Huang, "MSFT-YOLO: Improved YOLOv5 Based on Transformer for Detecting Defects of Steel Surface," *Sensors*, vol. 22, no. 9, p. 3467, 2022.

[12] Y. Xu, "Defect Detection of Industrial Aluminum Sheets Based on the Yolo Model," *Mathematical Modeling and Algorithm Applications*, vol. 7, pp. 38–48, 2025.

[13] X. Yu, B. Hu, W. Jiang, J. Wan, X. Yang, N. Liu, and X. Dong, "Enhanced YOLOv8 for industrial polymer films: a semi-supervised framework for micron-scale defect detection," *Frontiers in Artificial Intelligence*, vol. 8, pp. 1–17, 2025.

[14] B. H. Lin, J. C. Chen, and J. J. J. Lien, "Defect Inspection Using Modified YoloV4 on a Stitched Image of a Spinning Tool," *Sensors*, vol. 23, 2023.

[15] W. Liu, M. Jia, S. Zhang, S. Zhu, J. Qi, and J. Hu, "A Lightweight TA-YOLOv8 Method for the Spot Weld Surface Anomaly Detection of Body in White," *Applied Sciences*, vol. 15, no. 6, p. 2931, 2025.

[16] L. Arias-Martínez, F. Jáñez-Martino, E. Fidalgo, P. Rodríguez-González, A. I. Fernández-Abia, E. Alegre, and J. Barreiro, "Automated shrinkage and gas porosity detection using YOLO model in additive-manufactured aluminum alloy parts," *Integrating Materials and Manufacturing Innovation*, vol. 14, pp. 320–332, 2025.

[17] Z. Yuan, X. Tang, H. Ning, and Z. Yang, "LW-YOLO: Lightweight deep learning model for fast and precise defect detection in printed circuit boards," *Symmetry*, vol. 16, no. 4, p. 418, 2024.

[18] Z. Zhang, C. Niu, Z. Zhao, X. Zhang, and X. Chen, "Small object few-shot segmentation for vision-based industrial inspection," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 4, pp. 4388–4399, 2025.

[19] D. Im, S. Lee, H. Lee, B. Yoon, F. So, and J. Jeong, "A data-centric approach to design and analysis of a surface-inspection system based on deep learning in the plastic injection molding industry," *Processes*, vol. 9, no. 11, p. 1895, 2021.

[20] M. T. Okano, W. A. C. Lopes, S. M. Ruggero, O. Vendrametto, and J. C. L. Fernandes, "Edge AI for industrial visual inspection: YOLOv8-based visual conformity detection using Raspberry Pi," *Algorithms*, vol. 18, no. 8, p. 510, 2025.

[21] G. Guiqiang, J. Junbao, L. Chengzhang, and L. Shuo, "Edge-YOLO: Lightweight multi-scale feature extraction for industrial surface inspection," *IEEE Access*, vol. 13, pp. 48188–48201, 2025.

[22] A. Medina, R. Bradley, W. Xu, P. Ponce, B. Anthony, and A. Molina, "Learning manufacturing computer vision systems using tiny YOLOv4," *Frontiers in Robotics and AI*, vol. 11, p. 1331249, 2024.

[23] L. Yu, J. Zhu, Q. Zhao, and Z. Wang, "An efficient YOLO algorithm with an attention mechanism for vision-based defect inspection deployed on FPGA," *Micromachines*, vol. 13, no. 7, p. 1058, 2022.

[24] J. Kang, Y. Cen, Y. Cen, K. Wang, and Y. Liu, "CFIS-YOLO: A lightweight multi-scale fusion network for edge-deployable wood defect detection," *Wood Material Science and Engineering*, 2025.

[25] S. H. Park, K. H. Lee, J. S. Park, and Y. S. Shin, "Deep learning-based defect detection for sustainable smart manufacturing," *Sustainability*, vol. 14, no. 5, p. 2697, 2022.

[26] S. Y. Lin and H. Y. Li, "Integrated circuit board object detection and image augmentation fusion model based on YOLO," *Frontiers in Neurorobotics*, vol. 15, p. 762702, 2021.

[27] P. T. Nguyen, D. C. Huynh, L. D. Ho, and M. W. Dunnigan, "YOLO-WTB: Improved YOLOv12n model for detecting small damage of wind turbine blades from aerial imagery," *IEEE Access*, vol. 13, pp. 131257–131270, 2025.

[28] L. Li, A. Cherouat, H. Snoussi, T. Wang, Y. Lou, and Y. Wu, "Vision-based deep learning for robot grasping application in Industry 4.0," HAL preprint, 2024.

[29] A. S. Ludwika and A. P. Rifai, "Deep learning for detection of proper utilization and adequacy of personal protective equipment in manufacturing teaching laboratories," *Safety*, vol. 10, no. 1, p. 26, 2024.

[30] I. V. Fanthony, Z. Husin, H. Hikmarika, S. Dwijayanti, and B. Y. Suprapto, "YOLO algorithm-based surrounding object identification on autonomous electric vehicle," in *Proceedings of the 2021 EECSI*, 2021, pp. 1–6.

[31] M.H.F. Md Fauadi, M.F.H. Mohd Zan, M.A.M. Ali, L. Abdullah, S.N. Yaakop, and A.Z.M. Noor, "Enhancing Driving Safety and Environmental Consciousness through Automated Road Sign Recognition Using Convolutional Neural Networks", *Nature Environment and Pollution Technology,* vol. 23, no. 4 pp. 2461-2468, 2024.

[32] H. Xu, Z. Zhang, H. Ye, J. Song, and Y. Chen, "Efficient steel surface defect detection via a lightweight YOLO framework with task-specific knowledge-guided optimization," *Electronics*, vol. 14, no. 10, p. 2029, 2025.

[33] K. Ramesh, F. Mushtaq, S. Deshmukh, T. Ray, C. Parimi, A. Basem, and A. Elsheikh, "An investigation of deep learning approaches for efficient assembly component identification," *Beni-Suef University Journal of Basic and Applied Sciences*, vol. 13, p. 537, 2024.

[34] P. Dong, Y. Wang, Q. Yu, W. Feng, and G. Zong, "AMC-YOLO: Improved YOLOv8-based defect detection for cigarette packs," *IET Image Processing*, vol. 18, no. 4, p. 13272, 2024.

[35] C. Song, Y. Zhou, Y. Ma, Q. Qi, Z. Wang, and K. Hu, "YOLO-HDEW: An efficient PCB defect detection model," *Electronics*, vol. 14, no. 17, p. 3383, 2025.

[36] V. A. Adibhatla, H. C. Chih, C. C. Hsu, J. Cheng, M. F. Abbod, and J. S. Shieh, "Defect detection in printed circuit boards using you-only-look-once convolutional neural networks," *Electronics*, vol. 9, no. 9, p. 1547, 2020.

[37] Y. Wang, J. Huang, M. S. K. Dipu, H. Zhao, S. Gao, H. Zhang, and P. Lv, "YOLO-RLC: An advanced target-detection algorithm for surface defects of printed circuit boards based on YOLOv5," *Computers, Materials and Continua*, vol. 80, no. 2, pp. 4389–4405, 2024.

[38] M. Hussain, "YOLOv1 to v8: Unveiling each variant—A comprehensive review of YOLO," *IEEE Access*, vol. 12, p. 3378568, 2024.

[39] R. Sapkota, D. Ahmed, and M. Karkee, "Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments," *Artificial Intelligence in Agriculture*, vol. 13, p. 100247, 2024.

[40] A. S. Novodranov and E. V. Shapovalov, "Applying neural network technologies for quality control in robotic surfacing," in *Welding and Related Technologies – Proceedings of the 7th International Conference on Welding and Related Technologies (WRT 2024)*, 2025, pp. 247–252.

[41] S. J. Ji, Q. H. Ling, and F. Han, "An improved algorithm for small object detection based on YOLO v4 and multi-scale contextual information," *Computers and Electrical Engineering*, vol. 105, p. 108490, 2023.

[42] M. Hu, Z. Li, J. Yu, X. Wan, H. Tan, and Z. Lin, "Efficient-lightweight YOLO: Improving small object detection in YOLO for aerial images," *Sensors*, vol. 23, no. 14, p. 6423, 2023.

[43] M. Hussain, "YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection," *Machines*, vol. 11, no. 7, p. 677, 2023.

[44] M. Mao and M. Hong, "YOLO object detection for real-time fabric defect inspection in the textile industry: A review of YOLOv1 to YOLOv11," *Sensors*, vol. 25, no. 7, p. 2270, 2025.

[45] R. Sharma, B. B. Sharma, P. Vaidya, and S. Naaz, "Design and implementation of real-time tomato plant growth monitoring system using deep learning based YOLO and Raspberry Pi," *Applied Artificial Intelligence*, vol. 39, no. 1, p. 2590829, 2025.

[46] T. H. Khan, M. Manekar, N. Ramkelkar, V. Karadbhajne, S. Bhagat, and P. Kohaley, "Automated orange detection using YOLO on Raspberry Pi," *International Research Journal of Innovative Engineering and Technology*, vol. 8, 2024.

[47] U. Obu, Y. Ambekar, H. Dhote, S. Wadbudhe, S. Khandelwal, and S. Dongre, "Crop disease detection using YOLOv5 on Raspberry Pi," in *Proceedings of the 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN 2023)*, 2023, pp. 1–6.

[48] R. Al Amin, M. Hasan, V. Wiese, and R. Obermaisser, "FPGA-based real-time object detection and classification system using YOLO for edge computing," *IEEE Access*, vol. 12, p. 3404623, 2024.

[49] W. Weng, Z. Lai, Z. Cui, Z. Chen, H. Chen, T. Lin, J. Wang, S. Zheng, and G. Chen, "GCD-YOLO: A deep learning network for accurate tomato fruit stalks identification in unstructured environments," *Smart Agricultural Technology*, vol. 12, p. 101465, 2025.

[50] M. Kulshreshtha, S. S. Chandra, P. Randhawa, G. Tsaramirsis, A. Khadidos, and A. O. Khadidos, “OATCR: Outdoor autonomous trash-collecting robot design using YOLOv4-tiny,” *Electronics*, vol. 10, no. 18, p. 2292, 2021.

[51] C. Mwitta, G. C. Rains, and E. Prostko, “Evaluation of inference performance of deep learning models for real-time weed detection in an embedded computer,” *Sensors*, vol. 24, no. 2, p. 514, 2024.

[52] L. Y. Lo, C. H. Yiu, Y. Tang, A. S. Yang, B. Li, and C. Y. Wen, “Dynamic object tracking on autonomous UAV system for surveillance applications,” *Sensors*, vol. 21, no. 23, p. 7888, 2021.

[53] H. U. Rahman, M. Raza, P. Afsar, H. U. Khan, and S. Nazir, “Analyzing factors that influence offshore outsourcing decision of application maintenance,” *IEEE Access*, vol. 8, pp. 180913–180925, 2020.

[54] H. Jahanshahi and Z. H. Zhu, “Review of machine learning in robotic grasping control in space application,” *Acta Astronautica*, 2024.

[55] K. S. Khor, C. Liu, and C. C. Cheah, “Robotic grasping of unknown objects based on deep learning-based feature detection,” *Sensors*, vol. 24, no. 15, p. 4861, 2024.

[56] M. Imam, K. Baina, Y. Tabii, E. Mostafa Ressami, Y. Adlaoui, I. Benzakour, F. Bourzeix, and E. Hassan Abdelwahed, “Ensuring miners’ safety in underground mines through edge computing: Real-time PPE compliance analysis based on pose estimation,” *IEEE Access*, vol. 12, p. 3470558, 2024.

[57] A. S. Ludwika and A. P. Rifai, “Deep learning for detection of proper utilization and adequacy of personal protective equipment in manufacturing teaching laboratories,” *Safety*, vol. 10, no. 1, p. 26, 2024.

[58] K. S. Kavitha, C. G. Mamatha, and T. V. Rohini, “Deep learning based product defect detection for sustainable smart manufacturing,” in *Proceedings of the 5th IEEE Global Conference for Advancement in Technology (GCAT 2024)*, 2024, pp. 1–6.

[59] 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,” *Computers and Electrical Engineering*, vol. 108, p. 108706, 2023.

[60] M. V. Kurkute and G. Krishnamoorthy, “Real-time IoT data analytics for smart manufacturing: Leveraging machine learning for predictive analytics and process optimization in industrial systems,” *Journal of Science and Technology*, vol. 5, 2024.

[61] L. Xu, X. Zhang, F. Ma, G. Chang, C. Zhang, J. Li, S. Wang, and Y. Huang, “Detecting defects in fused deposition modeling based on improved YOLO v4,” *Materials Research Express*, vol. 10, p. 095304, 2023.

[62] A. R. Sani, A. Zolfagharian, and A. Z. Kouzani, “Automated defects detection in extrusion 3D printing using YOLO models,” *Journal of Intelligent Manufacturing*, 2024.

[63] F. Xia, G. Nan, Z. Jia, and D. Wang, "Enhanced YOLO with FPGA hardware acceleration for aluminum sheet defect detection," *Future Generation Computer Systems*, vol. 176, p. 108189, 2026.

[64] Y. Xu, "Defect detection of industrial aluminum sheets based on the YOLO model," *Mathematical Modelling and Algorithm Applications*, vol. 7, no. 2, pp. 38-48, 2025.

[65] W. Dai, Y. Wang, X. Li, and Y. Wang, "YOLO aluminum profile surface defect detection system for FPGA deployment," *Journal of Electronic Measurement and Instrumentation*, vol. 38, no. 9, pp. 160-167, 2023.

[66] J. Li, J. Gu, Z. Huang, and J. Wen, "Application research of improved YOLO v3 algorithm in PCB electronic component detection," *Applied Sciences*, vol. 9, no. 18, p. 3750, 2019.

[67] M.B. Akhtar, "The use of a convolutional neural network in detecting soldering faults from a printed circuit board assembly," *HighTech and Innovation Journal*, vol. 3, no. 1, pp. 1–10, 2022.