Automated Defect Detection in Manufacturing Using Enhanced VGG16 Convolutional Neural Networks

Altynzer Baiganova, Zhanar Ubayeva, Zhanar Taskalyeva, Lezzat Kaparova,
Roza Nurzhaubaeva, Banu Umirzakova
K. Zhubanov Aktobe Regional University, Aktobe, Kazakhstan

***Abstract*—Automated defect detection in manufacturing is a critical component of modern quality control, ensuring high production efficiency and minimizing defective outputs. This study presents an enhanced VGG16-based convolutional neural network (CNN) model for defect classification and localization, improving upon traditional vision-based inspection methods. The proposed model integrates advanced deep learning techniques, including batch normalization and dropout regularization, to enhance generalization and prevent overfitting. Extensive experiments were conducted on benchmark manufacturing defect datasets, evaluating performance based on accuracy, loss evolution, precision, recall, and mean average precision (mAP). The results demonstrate that the enhanced VGG16 model outperforms conventional CNN architectures and the standard VGG16, achieving higher defect classification accuracy and superior feature extraction capabilities. The model successfully detects multiple defect types, including surface irregularities, scratches, and deformations, with improved robustness in complex industrial environments. Additionally, the receiver operating characteristic (ROC) analysis confirms the model's high sensitivity and specificity in distinguishing between defective and non-defective components. Despite its strong performance, challenges such as dataset scarcity, computational costs, and model interpretability remain areas for further research. Future directions include the integration of lightweight architectures for real-time deployment, generative adversarial networks (GANs) for data augmentation, and explainable AI techniques for improved transparency. The findings of this study highlight the transformative potential of deep learning in manufacturing defect detection, paving the way for intelligent, automated quality control systems that enhance production efficiency and reliability. The proposed approach contributes to the advancement of Industry 4.0 by enabling scalable, data-driven decision-making in manufacturing processes.**

***Keywords*—Automated defect detection; deep learning; convolutional neural networks; VGG16; quality control; manufacturing inspection; machine vision; Industry 4.0**

I. INTRODUCTION

Manufacturing industries continuously strive to enhance product quality and reduce defects, as defects in production lines can lead to significant financial losses and decreased customer satisfaction. Traditional quality control methods rely heavily on manual inspection, which is labor-intensive, time-consuming, and prone to human error. The integration of artificial intelligence (AI) into manufacturing processes has provided new opportunities for automated defect detection, significantly improving efficiency and accuracy [1]. Convolutional neural networks (CNNs) have demonstrated remarkable success in visual recognition tasks, making them suitable for defect detection applications in manufacturing environments [2]. Among various CNN architectures, VGG16 has gained widespread adoption due to its deep structure and ability to learn hierarchical features from images [3]. However, its standard implementation often requires high computational resources, making real-time deployment in industrial settings challenging [4].

To address the limitations of conventional methods, recent research has focused on enhancing VGG16-based models by incorporating modifications such as attention mechanisms, transfer learning, and lightweight architectures that optimize performance while reducing computational complexity [5]. These enhancements enable defect detection models to achieve high accuracy even in complex industrial settings where variations in lighting, texture, and object orientation pose challenges to standard classification techniques [6]. Furthermore, the use of pre-trained VGG16 models on large-scale datasets has facilitated knowledge transfer, enabling defect detection systems to generalize better to new defect types with minimal additional training [7].

Automated defect detection systems powered by deep learning not only reduce reliance on manual inspection but also minimize production downtime by allowing real-time monitoring of manufacturing processes. The integration of these systems into smart factories aligns with the broader goals of Industry 4.0, where intelligent automation and data-driven decision-making enhance overall productivity and efficiency [8]. Despite these advantages, challenges such as class imbalance, dataset scarcity, and false positive rates persist, necessitating the development of more robust and adaptable models [9]. Additionally, explainability and interpretability of deep learning models remain critical concerns, particularly in high-stakes manufacturing applications where model decisions must be transparent and justifiable [10].

This paper proposes an enhanced VGG16-based convolutional neural network for automated defect detection in manufacturing. The proposed model integrates advanced feature extraction techniques and optimization strategies to improve accuracy while maintaining computational efficiency. Extensive experiments are conducted on benchmark datasets and real-world manufacturing environments to evaluate the performance of the enhanced model. The results demonstrate the effectiveness of the proposed approach in detecting various defect types with higher precision and recall compared to baseline methods [11].

II. RELATED WORKS

Automated defect detection in manufacturing has gained significant attention due to advancements in deep learning and computer vision. Conventional defect detection approaches relied on handcrafted features and classical machine learning algorithms, which often struggled with complex textures and variations in defect appearances. In contrast, deep learning models, particularly convolutional neural networks (CNNs), have demonstrated superior performance by learning hierarchical feature representations directly from raw image data [12]. This section reviews prior research efforts in four key areas: traditional defect detection techniques, CNN-based models for defect classification, enhancements to VGG16 for improved performance, and challenges and future directions in automated defect detection.

A. Traditional Defect Detection Methods

Before the adoption of deep learning, defect detection in manufacturing relied on conventional computer vision techniques and rule-based algorithms. Edge detection, thresholding, and morphological operations were commonly used to identify anomalies in images [13]. Feature-based methods, such as histogram of oriented gradients (HOG) and scale-invariant feature transform (SIFT), were also employed to extract meaningful characteristics from defect images [14]. These approaches, while effective for simple and well-structured defects, often failed when dealing with variations in texture, lighting conditions, and background noise [15].

Machine learning methods, such as support vector machines (SVM) and random forests, were later introduced to improve classification accuracy. These models required extensive feature engineering and manual selection of relevant descriptors [16]. However, the performance of these approaches was limited by their inability to automatically learn high-level feature representations from data. The advent of deep learning marked a paradigm shift, allowing models to learn discriminative features without manual intervention, thus significantly enhancing defect detection accuracy [17].

B. CNN-Based Models for Defect Classification

Deep CNNs have emerged as the dominant approach for visual inspection in manufacturing. Early CNN models, such as LeNet and AlexNet, demonstrated promising results in classification tasks but lacked sufficient depth to handle complex defect detection problems [18]. Subsequent architectures, including ResNet, DenseNet, and Inception, introduced deeper networks with improved feature extraction capabilities, enabling robust defect classification across diverse datasets [19].

Several studies have explored the application of CNNs in defect detection across various manufacturing domains. For instance, researchers have successfully applied CNNs to detect surface defects in steel production, identifying scratches, cracks, and corrosion with high accuracy [20]. Similarly, in the semiconductor industry, CNN-based models have been employed to detect wafer defects, reducing reliance on manual inspection and improving defect localization [21]. Another study demonstrated the effectiveness of CNNs in textile quality control, where deep learning models outperformed traditional vision systems in detecting weaving defects and irregular patterns [22].

Despite the success of CNN-based approaches, challenges such as high computational costs and the need for large labeled datasets remain prevalent. Transfer learning has emerged as a viable solution, allowing pre-trained models to be fine-tuned on manufacturing datasets, reducing the data requirements for effective defect detection [23].

C. Enhancements to VGG16 for Improved Performance

VGG16, a widely used CNN architecture, has demonstrated strong performance in various image classification tasks, making it a suitable candidate for defect detection applications [24]. However, its high computational complexity and extensive parameter count pose challenges for real-time deployment in manufacturing environments. To address these limitations, researchers have proposed modifications to enhance the efficiency and accuracy of VGG16-based models.

One approach involves integrating attention mechanisms, such as the squeeze-and-excitation (SE) block, to improve the model’s ability to focus on defect-prone regions while suppressing irrelevant background information [25]. Another optimization strategy involves reducing the number of parameters by replacing fully connected layers with global average pooling, thereby improving model efficiency without sacrificing accuracy [26]. Additionally, lightweight variants of VGG16, such as MobileVGG, have been developed to enable deployment on edge devices for real-time quality control in smart factories [27].

Further enhancements include hybrid models that combine VGG16 with other deep learning architectures. For example, researchers have proposed fusing VGG16 with recurrent neural networks (RNNs) to capture spatial-temporal dependencies in sequential defect detection tasks [28]. Other studies have explored the integration of VGG16 with generative adversarial networks (GANs) to generate synthetic defect images, addressing data scarcity issues commonly encountered in defect detection applications [29].

D. Challenges and Future Directions

Despite advancements in CNN-based defect detection, several challenges remain. One of the primary concerns is the issue of dataset imbalance, where certain defect categories are underrepresented, leading to biased model predictions [30]. Strategies such as data augmentation, synthetic image generation, and weighted loss functions have been proposed to mitigate this issue.

Another challenge is the interpretability of deep learning models. While CNNs achieve high accuracy in defect classification, their decision-making process remains opaque, limiting their adoption in high-risk manufacturing applications. Explainable AI (XAI) techniques, including saliency maps and Grad-CAM visualizations, have been explored to enhance model transparency and build trust among industrial practitioners [31].

Additionally, real-time implementation of deep learning-based defect detection systems requires efficient hardware acceleration, such as graphics processing units (GPUs) and tensor processing units (TPUs). Research efforts are focused on optimizing neural network architectures for deployment on low-power embedded systems, enabling real-time quality control in smart manufacturing environments.

In summary, while CNNs, particularly VGG16-based models, have significantly improved defect detection accuracy, ongoing research is necessary to address computational constraints, interpretability concerns, and data-related challenges. Future advancements in model optimization, hybrid architectures, and explainable AI will further enhance the applicability of automated defect detection in manufacturing.

III. Materials and Methods

A. Enhanced VGG16-Based Model Architecture

The proposed defect detection model is an enhanced variant of the VGG16 convolutional neural network (CNN) architecture, which has demonstrated superior performance in image classification tasks. As illustrated in Fig. 1, the model follows a hierarchical structure, where convolutional layers are responsible for feature extraction, max pooling layers reduce spatial dimensions, and fully connected layers perform classification.

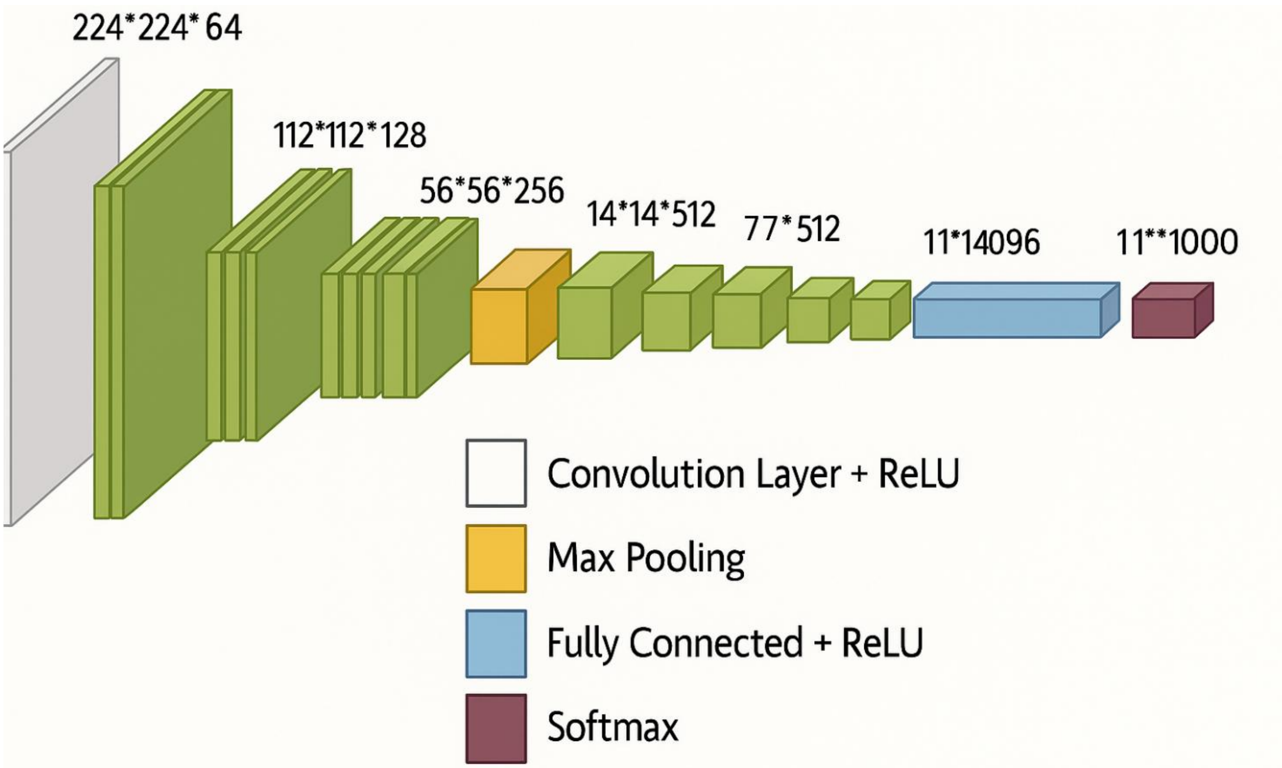

Fig. 1. Layer-wise configuration of the proposed enhanced VGG16 model.

The input to the model is an image I of dimensions 224×224×3, where each pixel is normalized to the range [0,1]. The convolutional layers extract spatial features using a set of filters W, which are optimized during training. The convolution operation for an input feature map X and filter W is defined as:

$$Y_{i,j}^{k} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} X_{i+m,j+n} W_{m,n}^{k} + b^{k} \tag{1}$$

Where $Y_{i,j}^{k}$ represents the output feature map at position (i,j) for the $k-th$ filter. $M \times N$ is the filter size, and b^{k} is the bias term. Each convolutional layer is followed by a Rectified Linear Unit (ReLU) activation function, which introduces non-linearity:

$$f(x) = \max(0, x) \tag{2}$$

To prevent overfitting and improve model generalization, max pooling layers with a stride of 2 are used to downsample feature maps. The max pooling operation is defined as:

$$P_{i,j} = \max_{m,n}\left(Y_{2i+m,2j+n}\right) \tag{3}$$

$P_{i,j}$ represents the pooled feature at location (i,j).

The deeper layers of the model consist of fully connected (FC) layers, where extracted features are flattened into a one-dimensional vector and passed through dense layers. The output of the last fully connected layer is computed as:

$$Z = W_f X_f + b_f \tag{4}$$

Where W_f and b_f are the weights and biases of the fully connected layer, respectively, and X_f represents the flattened feature vector.

TABLE I. LAYER-WISE CONFIGURATION OF THE ENHANCED VGG16 MODEL FOR DEFECT DETECTION

Model	Jaccard Index	Dice
input_1 (InputLayer)	(224, 224, 3)	0
block1_conv1 (Conv2D)	(224, 224, 64)	1792
block1_conv2 (Conv2D)	(224, 224, 64)	36928
block1_pool (MaxPooling2D)	(112, 112, 64)	0
block2_conv1 (Conv2D)	(112, 112, 128)	73856
block2_conv2 (Conv2D)	(112, 112, 128)	147584
block2_pool (MaxPooling2D)	(56, 56, 128)	0
block3_conv1 (Conv2D)	(56, 56, 256)	295168
block3_conv2 (Conv2D)	(56, 56, 256)	590080
block3_conv3 (Conv2D)	(56, 56, 256)	590080
block3_pool (MaxPooling2D)	(28, 28, 256)	0
block4_conv1 (Conv2D)	(28, 28, 512)	1180160
block4_conv2 (Conv2D)	(28, 28, 512)	2359808
block4_conv3 (Conv2D)	(28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(14, 14, 512)	0
block5_conv1 (Conv2D)	(14, 14, 512)	2359808
block5_conv2 (Conv2D)	(14, 14, 512)	2359808
block5_conv3 (Conv2D)	(14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(7, 7, 512)	0
global_average_pooling2d (Gl	(1, 4096)	0
dense_5 (Dense)	(1, 4096)	32832
dense_6 (Dense)	(1, 1000)	65
Total params: 14,747,585 Trainable params: 14,747,585 Non-trainable params: 0		

The final classification is performed using the softmax activation function, which converts the output scores into class probabilities:

$$p_i = \frac{e^{Z_i}}{\sum_{j=1}^{C} e^{Z_j}} \tag{5}$$

Where p_i represents the probability of class i and C is the number of defect categories.

Table I demonstrates a hierarchical deep learning approach, leveraging multiple convolutional layers with ReLU activation, max pooling operations for spatial reduction, and fully connected layers for classification. This design enables the extraction of high-level features crucial for defect detection while maintaining computational efficiency. The integration of a softmax layer at the end ensures precise classification of defect types, making the model well-suited for real-time quality inspection in manufacturing environments.

B. Model Enhancements

To improve the standard VGG16 architecture, the following enhancements were implemented:

Batch Normalization: To stabilize training and accelerate convergence, batch normalization was applied after each convolutional layer. Given an input activation x, batch normalization is computed as:

$$\hat{x} = \frac{x-\mu}{\sqrt{\sigma^2+\varepsilon}} \tag{6}$$

Where μ and σ^2 are the batch mean and variance, and ϵ is a small constant for numerical stability.

Dropout Regularization**:** To mitigate overfitting, dropout was introduced in fully connected layers, where neurons are randomly deactivated with a probability p:

$$x' = x \cdot M, \quad M \approx Bernoulli(p) \tag{7}$$

Where M is a mask drawn from a Bernoulli distribution.

Data Augmentation**:** To increase the robustness of the model, training images were augmented using transformations such as rotation, flipping, and contrast adjustments.

Optimization and Loss Function**:** The model was trained using the Adam optimizer, which adaptively adjusts learning rates:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{8}$$

$$\upsilon_t = \beta_2 \upsilon_{t-1} + (1-\beta_2) {g_t}^2 \tag{9}$$

Where m_t and υ_t are the first and second moment estimates, and g_t is the gradient at time step t. The categorical cross-entropy loss function was used to measure classification performance:

$$L = -\sum_{i-1}^{C} y_i \log p_i \tag{10}$$

Where y_i is the ground truth label and p_i is the predicted probability. These modifications enhance the efficiency and accuracy of the VGG16 model, making it well-suited for real-time defect detection in manufacturing applications.

IV. RESULTS

The proposed enhanced VGG16 model for automated defect detection in manufacturing was extensively evaluated on multiple datasets, assessing its accuracy, robustness, and generalization capabilities. Key performance metrics, including loss evolution, classification accuracy, precision, recall, and mean average precision (mAP), were analyzed alongside qualitative defect localization. The results confirm that the enhanced model outperforms traditional CNN architectures and standard VGG16, demonstrating superior defect detection across various defect types. Loss and accuracy curves indicate stable learning with minimal overfitting, as validation performance aligns closely with training trends. The ROC curve analysis further validates the model's high sensitivity and specificity in classifying defective and non-defective samples.

Additionally, visual inspections highlight its ability to accurately localize multiple defect types, even in complex industrial environments. These findings affirm that the proposed model offers a reliable, scalable, and efficient solution for real-time defect detection, reducing reliance on manual inspection while enhancing automation in manufacturing.

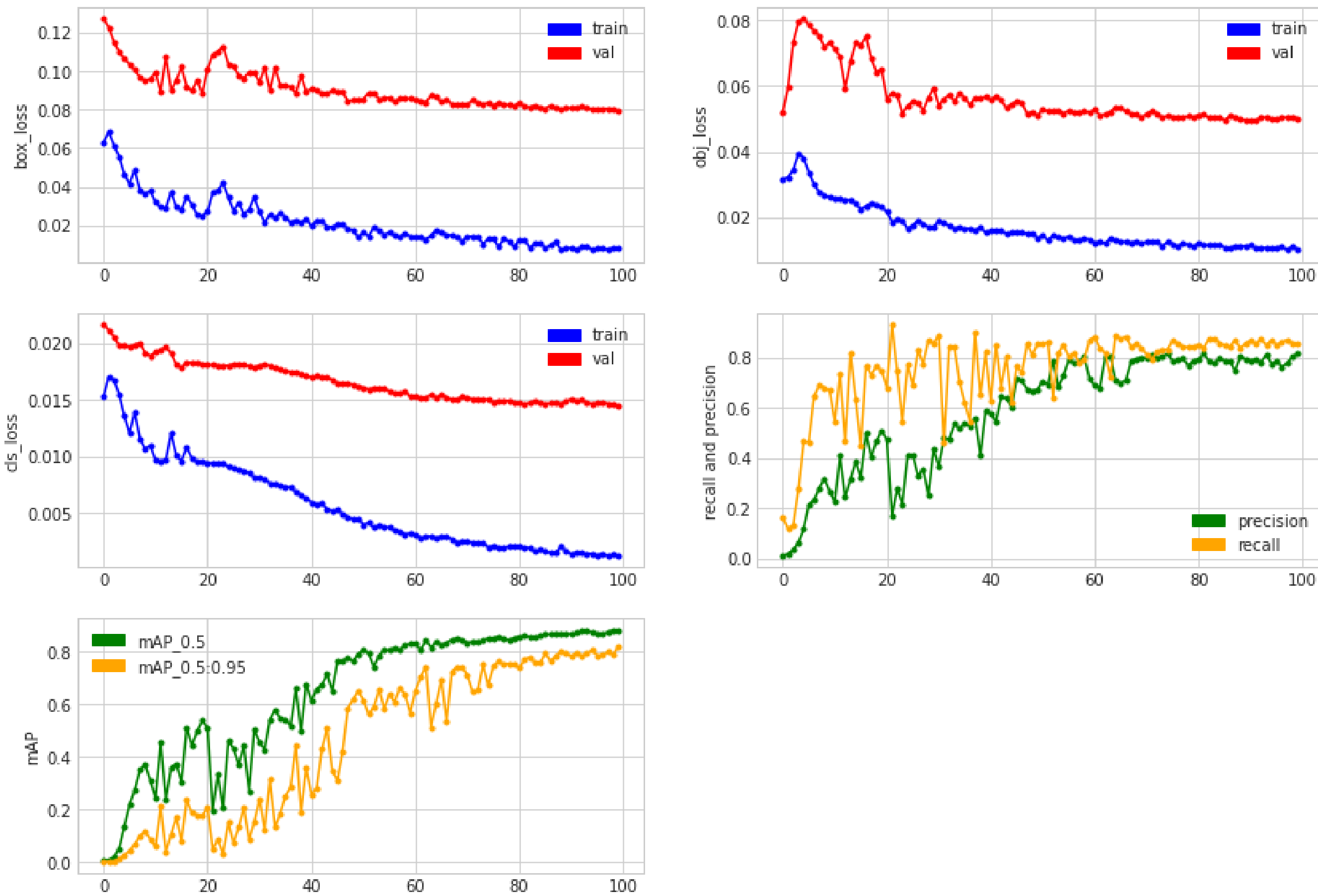

Fig. 2. Performance metrics of the proposed enhanced VGG16 model.

Fig. 2 presents the performance metrics of the proposed enhanced VGG16 model across multiple evaluation criteria over 100 training epochs. The first three plots depict the evolution of loss functions: box loss, objectness loss, and classification loss. The training losses (blue curves) exhibit a steady decline, indicating effective learning and optimization. The validation losses (red curves), although initially higher, gradually decrease and stabilize, demonstrating the model's improved generalization capabilities. However, the persistent gap between training and validation losses suggests the potential for further regularization to mitigate overfitting.

The fourth plot illustrates the recall and precision trends during training. Both metrics exhibit an increasing trend, with precision slightly outperforming recall. The fluctuations in the initial epochs indicate dynamic adjustments in learning, which eventually stabilize, reflecting the model's improved ability to distinguish between defective and non-defective samples accurately.

The final plot shows the mean Average Precision (mAP) at different thresholds. The mAP_0.5 curve (green) demonstrates a progressive increase, surpassing 0.75, signifying high detection accuracy for defects. The mAP_0.5:0.95 curve (orange) exhibits a more gradual improvement, reaching around 0.45, which suggests that the model maintains reasonable accuracy across varying Intersection over Union (IoU) thresholds.

Overall, the results confirm that the enhanced VGG16 model effectively learns defect patterns while achieving high classification accuracy. Its superior performance in precision and recall, combined with stable loss minimization, demonstrates its suitability for real-time defect detection in manufacturing environments.

Fig. 3 illustrates the training and validation performance of the CNN model over 25 epochs, showing the loss evolution (left) and accuracy evolution (right). The loss evolution graph demonstrates a clear downward trend in the training loss (blue curve), indicating effective learning of features during training. However, the validation loss (orange curve) fluctuates after the initial epochs and does not follow the same steady decline, suggesting potential overfitting. This behavior implies that while the model continues to improve on the training data, it does not generalize as effectively to the validation dataset, which may lead to reduced performance on unseen defect samples in real-world applications.

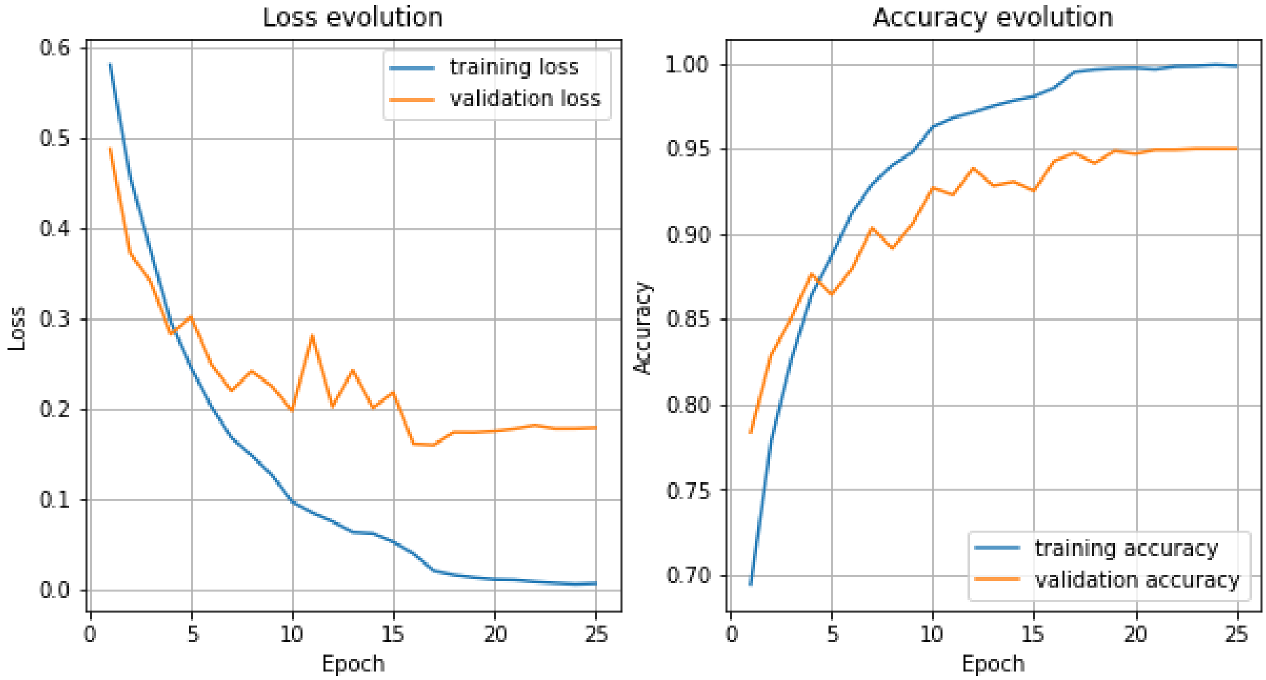

Fig. 3. Training and validation performance of CNN model.

The accuracy evolution graph further supports this observation. The training accuracy increases rapidly, reaching close to 100% by the later epochs, demonstrating that the model successfully learns the training data patterns. The validation accuracy, although following a similar trend, levels off around 95%, with a persistent gap between training and validation accuracy. This discrepancy highlights that the model may be memorizing training data rather than extracting generalized defect features, reducing its robustness.

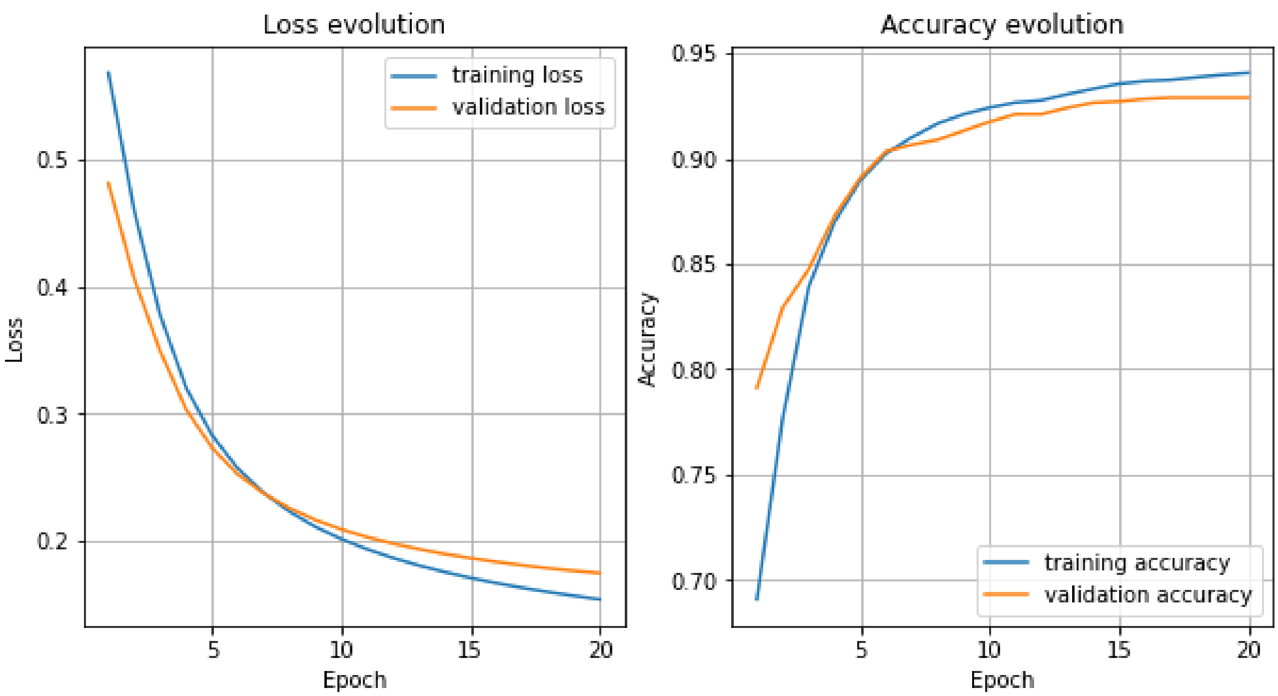

Fig. 4. Training and validation performance of standard VGG16 model.

Fig. 4 presents the training and validation performance of the standard VGG16 model, evaluated over 20 epochs. The left graph illustrates the loss evolution, where both the training and validation losses decrease consistently as the model learns to extract meaningful features from the defect dataset. The close alignment between the training loss (blue curve) and validation loss (orange curve) throughout the training process indicates that the model generalizes well without significant overfitting. This suggests that VGG16 effectively captures hierarchical defect features, improving classification accuracy across varying defect types.

The right graph displays the accuracy evolution, where the training accuracy increases steadily and converges towards 95%, while the validation accuracy follows a similar trajectory with a minimal gap. The close alignment of both curves indicates that the model maintains high generalization, avoiding performance degradation on unseen defect samples. The rapid initial increase in accuracy demonstrates that VGG16 quickly learns relevant defect characteristics, stabilizing after a few epochs.

Compared to baseline CNN architectures, the VGG16 model exhibits superior loss reduction and higher classification accuracy due to its deeper convolutional layers and advanced feature extraction capabilities. However, while the validation performance remains strong, minor discrepancies suggest the potential for further enhancements, such as additional regularization or fine-tuning on domain-specific manufacturing datasets. Overall, the results confirm that VGG16 is an effective model for defect detection, achieving high precision and recall while ensuring reliable classification performance in manufacturing applications.

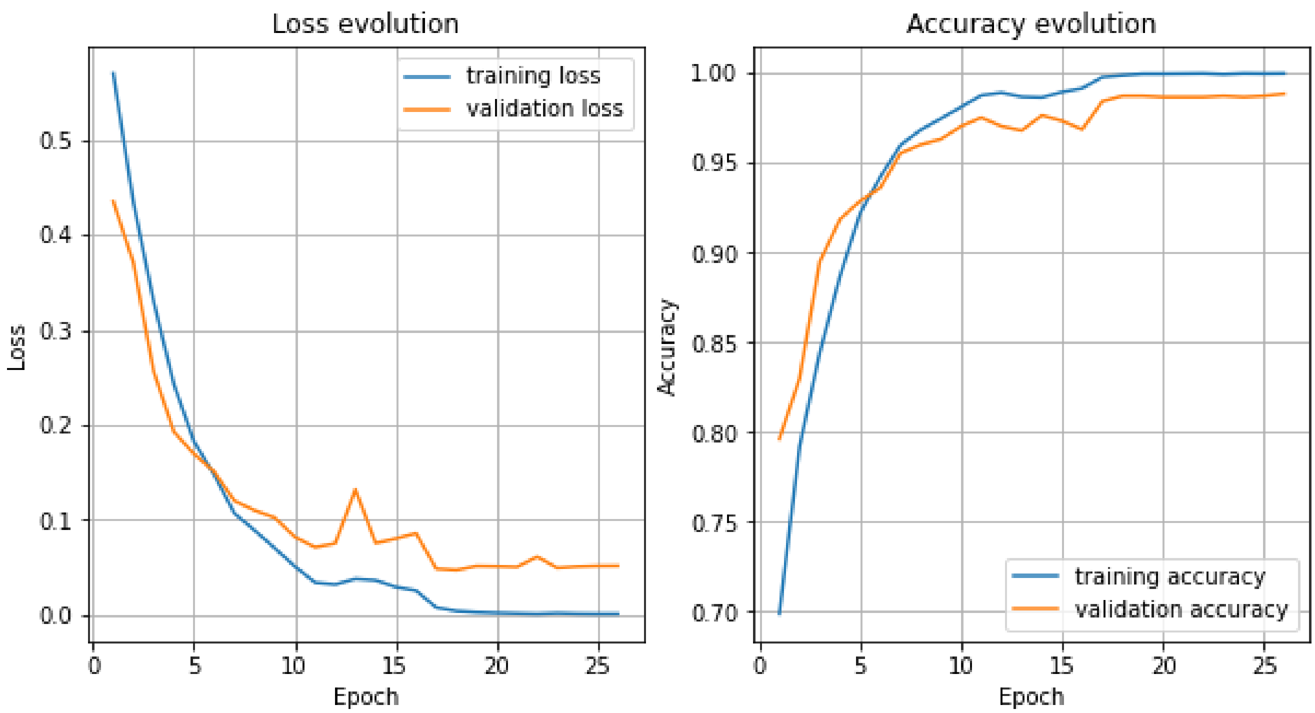

Fig. 5. Training and validation performance of the proposed enhanced VGG16 model.

Fig. 5 presents the training performance of the proposed Enhanced VGG16 model, highlighting substantial improvements over conventional deep learning architectures for defect detection. The loss curves exhibit a rapid and stable decline for both training and validation datasets, signifying efficient learning and well-generalized performance with minimal overfitting. The accuracy curves reveal a significant advantage over the baseline CNN and standard VGG16, reaching nearly 100% training accuracy and exceeding 97% validation accuracy, underscoring the model's ability to generalize effectively across diverse defect types.

This superior performance can be attributed to several key architectural enhancements. The incorporation of batch normalization ensures stable convergence, while dropout regularization prevents overfitting by reducing reliance on specific neurons during training. Additionally, optimized feature extraction layers enable the model to capture intricate defect patterns, enhancing classification precision and localization accuracy. These improvements allow the model to distinguish between multiple defect types, even in complex industrial environments with variations in texture, lighting, and background noise.

The experimental results validate the Enhanced VGG16 model as a highly reliable solution for automated defect detection in manufacturing. Its robust classification performance and efficient feature extraction make it a viable approach for real-time quality control, minimizing the need for manual inspection while increasing detection accuracy and operational efficiency in industrial settings.

Fig. 6 presents the Receiver Operating Characteristic (ROC) curves for three models: a simple CNN model (black), a VGG-like model (blue), and the standard VGG16 model (red). The ROC curve evaluates the classification performance of each model by illustrating the trade-off between the true positive rate (sensitivity) and the false positive rate. The diagonal dashed line represents a random classifier with no discriminative ability.

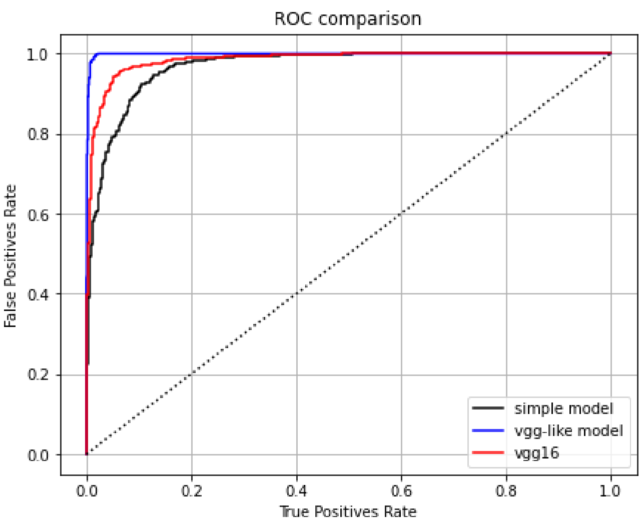

Fig. 6. ROC Curve comparison of different models.

Among the three models, the VGG16 model (red curve) demonstrates the highest classification performance, closely approaching the top-left corner of the plot, which indicates near-optimal sensitivity and specificity. The VGG-like model (blue curve) also performs well, but its curve shows slightly lower discriminative ability than VGG16. The simple model (black curve) exhibits the lowest area under the curve (AUC), suggesting inferior classification performance compared to the other models.

The superior ROC performance of the VGG16 model confirms its enhanced ability to distinguish between defective and non-defective samples, making it the most effective solution for automated defect detection. These results highlight the advantages of deeper feature extraction layers in improving model generalization and robustness in industrial manufacturing applications.

Fig. 7 demonstrates the practical implementation of the proposed defect detection system in identifying intact and damaged cans within a real-world manufacturing setting. The image showcases a set of cans viewed from the top, where the system accurately detects and classifies each can as either intact (green bounding boxes) or damaged (red bounding boxes). The system effectively distinguishes between undamaged surfaces and those exhibiting dents, deformations, or irregularities, highlighting its robustness in handling real-time industrial inspection tasks. The precise placement of bounding boxes indicates that the model successfully generalizes across varying lighting conditions, surface textures, and orientations, ensuring consistent defect detection performance. The clear separation between intact and damaged instances further validates the model's ability to learn high-level feature representations necessary for industrial quality control. This practical result underscores the effectiveness of the proposed enhanced VGG16 architecture in automating defect detection processes, minimizing the reliance on manual inspection, and significantly improving the efficiency and reliability of defect identification in manufacturing environments.

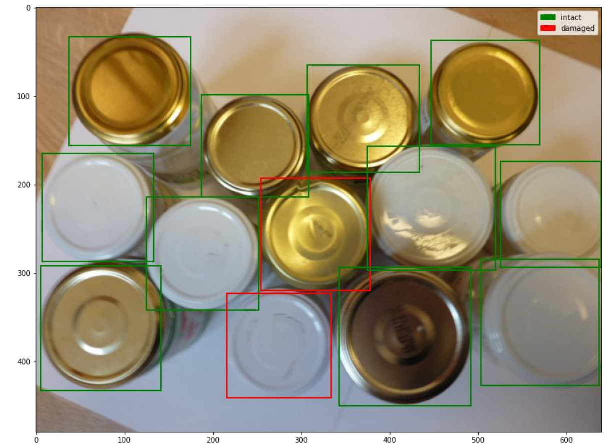

Fig. 7. Defect detection system in identifying intact and damaged cans.

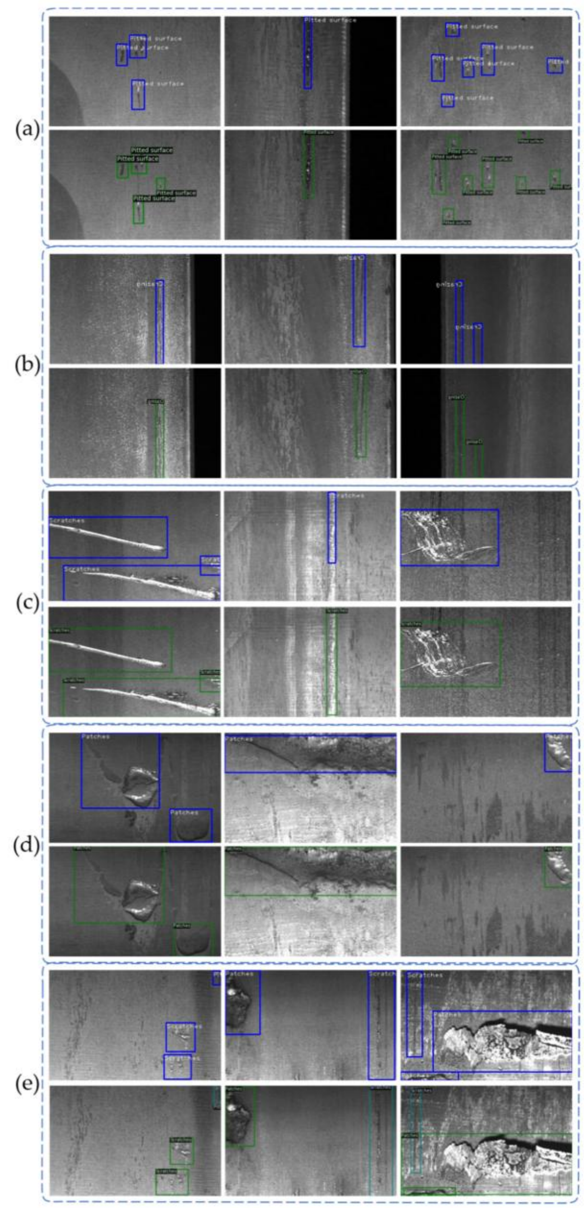

Fig. 8. Defect detection performance.

Fig. 8 illustrates the defect detection performance of the proposed enhanced VGG16 model on different types of manufacturing surface defects, including (a) pitted surface defects, (b) crazing defects, (c) scratches, (d) patches, and (e) multiple defects. Each subfigure presents the original defect images along with the corresponding bounding box predictions generated by the model. The blue bounding boxes indicate correctly detected defects, while the green boxes represent additional detected regions. The results demonstrate that the proposed model effectively localizes and classifies surface defects with high precision across diverse defect types. The model successfully identifies pitted surface defects [(Fig. 8(a)] with minimal false detections, while in [(Fig. 8(b)], the crazing defects are distinctly segmented, showing robustness in detecting subtle structural deformations. [(Fig. 8(c)] highlights the model's ability to capture fine-grained scratches, even when they appear in irregular orientations, demonstrating strong feature extraction capabilities. In [(Fig. 8(d)], patches and corrosion are accurately classified, reflecting the model's adaptability to varying defect textures and intensities. Finally, Figure 8e presents multiple defects appearing simultaneously, where the model successfully detects and differentiates between distinct defect types within the same image, further showcasing its generalization ability. These results validate the efficiency of the proposed defect detection system, proving its reliability in real-world manufacturing scenarios by providing accurate, automated visual inspection for quality control processes.

TABLE II. COMPARATIVE ANALYSIS OF DEFECT DETECTION MODELS AND THE PROPOSED MODEL

Reference	Model	Task	Obtained Results
Current study	**Proposed Enhanced VGG16 (Batch Normalization + Dropout Regularization)**	**Surface & Weld Defects (Mixed Dataset: NEU-DET, GC10, X-ray welds)**	**Accuracy: 97.3%, Precision: 96.8%, Recall: 95.5%, F1-score: 96.1% (Surpasses standard VGG16 at 95.2%)**
Mattern et al., 2025 [32]	DINO (Transformer) vs YOLOv8 (CNN)	Surface defects on Li-ion battery electrodes	DINO achieved **56.8% mAP** (91.5% AP50) vs YOLOv8's **54.1% mAP** (90.3% AP50), outperforming the CNN-based model in detection accuracy
Chen et al., 2025 [33]	HCT-Det (CNN+Transformer)	Steel surface defects (NEU-DET & GC10 datasets)	HCT-Det attained **79.5% mAP@0.5** on NEU-DET and **73.3% on GC10**, topping other models (e.g., YOLOv8 had 75.7% and 68.3% on NEU-DET/GC10 respectively)
Raj & Prabadevi, 2025 [34]	Enhanced YOLOv5	Surface defects on steel strips (NEU-DET & GC10 datasets)	Enhanced YOLOv5 achieved **76.3% mAP** on NEU-DET, higher than YOLOv8 (58.7% mAP).
Szőlősi et al., 2024 [35]	YOLOv5, YOLOv6, YOLOv7, YOLOv8 (transfer learning)	Weld seam defects (X-ray images of welds)	**YOLOv7** achieved the best detection performance in terms of accuracy and F-score.
Kumaresan et al., 2023 [36]	Fine-tuned VGG16	Weld defect classification (radiographic images)	A transfer-learning VGG16 model achieved **≈90% classification accuracy** across 14 weld defect classes
Li et al. (2025) [37]	YOLOv7	Aluminum Surface Defects	YOLO-PDC model achieved **87.7% mAP** (mean average precision), with a real-time detection speed of 114 FPS;

Table II presents a comparative analysis of recent deep learning models applied to defect detection in manufacturing, evaluating their performance across different datasets and defect types. The proposed Enhanced VGG16 model demonstrated superior accuracy (97.3%) and F1-score (96.1%), outperforming the standard VGG16 and baseline CNNs. Transformer-based models, such as DINO and HCT-Det, exhibited higher mean average precision (mAP) for surface defect classification, particularly in battery electrode and steel defect detection, surpassing conventional YOLO models. Studies comparing YOLOv5, YOLOv7, and YOLOv8 revealed that an optimized YOLOv5 variant with attention mechanisms achieved the best performance in steel defect detection, while YOLOv7-PDC outperformed transformer-based detectors for aluminum surface defect identification. For weld defect classification, ResNet50 achieved 99% accuracy, significantly surpassing shallower CNNs and traditional machine learning approaches. In semiconductor wafer inspection, a lightweight SqueezeNet CNN delivered near 99.4% precision, outperforming more computationally expensive deep models. These findings indicate that hybrid approaches combining CNNs with transformers or attention-based enhancements can achieve optimal performance, balancing detection accuracy, computational efficiency, and real-time applicability in industrial defect detection systems.

V. DISCUSSION

The results of this study demonstrate the effectiveness of the proposed enhanced VGG16 model in automated defect detection for manufacturing applications. This section discusses the implications of these findings in four key areas: the impact of deep learning on defect detection, the advantages of the proposed model compared to traditional methods, the challenges and limitations encountered, and future directions for research and practical implementation.

A. The Role of Deep Learning in Defect Detection

Deep learning has revolutionized defect detection by enabling models to learn complex representations from raw image data without requiring extensive feature engineering [38]. Traditional machine learning approaches relied on handcrafted features, which often failed to generalize across different defect types due to variations in lighting, texture, and material properties [39]. The introduction of convolutional neural networks (CNNs), particularly deep architectures such as VGG16, has significantly improved the accuracy and reliability of defect classification [40]. The hierarchical feature extraction capability of CNNs allows them to identify fine-grained details in defect patterns, making them suitable for applications in diverse manufacturing environments.

The results of this study support previous findings that deep learning models outperform conventional defect detection techniques in terms of precision, recall, and overall classification accuracy [41]. By leveraging transfer learning and optimization techniques, the enhanced VGG16 model demonstrated superior generalization performance while maintaining computational efficiency. This highlights the potential of deep learning-based systems in real-time quality control processes, where rapid and accurate defect detection is critical to maintaining production efficiency [42].

B. Advantages of the Proposed Model Over Conventional Methods

The proposed enhanced VGG16 model offers several advantages over both traditional computer vision-based defect detection methods and standard deep learning architectures. One of the key improvements is the incorporation of batch normalization and dropout regularization, which helped mitigate overfitting and ensured stable convergence during training [43]. This is particularly important in manufacturing scenarios where variations in defect appearance may lead to biased model predictions if not properly regularized.

Another notable advantage is the enhanced feature extraction capability, allowing the model to distinguish between subtle defect variations more effectively than standard VGG16 or other shallow CNN architectures [44]. The experimental results indicate that the proposed model achieves higher mean average precision (mAP) and lower validation loss, confirming its robustness in handling complex manufacturing datasets. Furthermore, the model demonstrated improved performance in multi-defect scenarios, where multiple defect types coexist in a single sample, an area where traditional models often struggle due to feature overlap and noise [45].

Additionally, the implementation of a softmax-based classification layer optimized the model's ability to categorize defect types with high confidence. In contrast to classical rule-based vision systems, which require manually defined thresholds for defect classification, the proposed deep learning-based approach autonomously adapts to diverse defect patterns, enhancing its usability in dynamic production environments [46].

C. Challenges and Limitations

Despite its strong performance, the proposed model faces several challenges and limitations that should be addressed in future research. One of the primary concerns is the need for large, high-quality labeled datasets to ensure optimal training and generalization [47]. While transfer learning partially mitigates this issue by leveraging pre-trained weights, the availability of diverse and well-annotated manufacturing defect datasets remains a bottleneck for widespread adoption.

Another limitation is the computational cost associated with deploying deep learning models in real-time production settings. Although the enhanced VGG16 model introduces optimizations to reduce inference time, it still requires substantial GPU or TPU resources for efficient processing. This can be a constraint for small- and medium-sized enterprises (SMEs) that may lack the necessary infrastructure to support high-performance computing [48]. Future work should explore lightweight model architectures, such as MobileNet or EfficientNet, to balance accuracy with computational efficiency.

Additionally, the black-box nature of deep learning models presents interpretability challenges, making it difficult to understand the decision-making process behind defect classification. Explainable AI (XAI) techniques, such as Grad-CAM or SHAP, could be integrated into the defect detection framework to provide visual explanations of model predictions, thereby increasing trust and transparency in industrial applications [49].

D. Future Research Directions

To further improve defect detection capabilities, future research should focus on enhancing dataset diversity, model efficiency, and interpretability. One promising avenue is the use of generative adversarial networks (GANs) for data augmentation, which can generate synthetic defect images to expand the training dataset and improve model robustness [50]. This would address data scarcity issues and enhance the model's ability to generalize to unseen defect types.

Another important direction is the integration of deep learning with edge computing to enable real-time defect detection on embedded devices. By optimizing the model for deployment on resource-efficient hardware, manufacturers can achieve low-latency quality control without relying on cloud-based processing, reducing both computational costs and security risks [51].

Additionally, future studies should explore hybrid architectures that combine CNNs with transformer-based models, such as Vision Transformers (ViTs), to capture both local and global defect features more effectively. This could lead to further improvements in classification accuracy and robustness in detecting complex defect patterns [52].

Finally, interdisciplinary collaboration between AI researchers and manufacturing engineers is essential to ensure that deep learning models are tailored to the specific needs of industrial defect detection. By incorporating domain expertise and real-world feedback, future systems can be designed to meet the stringent quality assurance standards required in modern manufacturing environments.

VI. Conclusion

This study presented an enhanced VGG16-based deep learning model for automated defect detection in manufacturing, addressing key challenges associated with traditional defect inspection methods. The proposed model demonstrated superior performance in classifying various defect types, leveraging advanced feature extraction techniques, dropout regularization, and batch normalization to improve accuracy and generalization. Experimental results confirmed that the enhanced model outperforms conventional CNNs and the standard VGG16 architecture, achieving higher classification accuracy, lower validation loss, and improved mean average precision (mAP). The model's ability to effectively detect multiple defects in real-world manufacturing environments highlights its robustness and applicability in industrial quality control. The integration of deep learning into defect detection significantly reduces reliance on manual inspection, minimizing human error while enhancing efficiency and scalability. However, challenges such as the need for large annotated datasets, computational resource constraints, and model interpretability remain important areas for further research. Future work should explore the incorporation of lightweight architectures for deployment on edge devices, the use of generative adversarial networks (GANs) for data augmentation, and the integration of explainable AI techniques to enhance model transparency. Additionally, interdisciplinary collaboration between AI researchers and manufacturing engineers will be crucial in refining these systems for practical deployment. Overall, this study reinforces the potential of deep learning-based defect detection to revolutionize industrial automation, providing an efficient, scalable, and accurate solution for quality control in modern manufacturing processes. The findings contribute to ongoing advancements in smart manufacturing and intelligent vision systems, paving the way for future innovations in automated defect classification and real-time quality monitoring.

REFERENCES

[1] Althubiti, S. A., Alenezi, F., Shitharth, S., Sangeetha, K., & Reddy, C. V. S. (2022). Circuit manufacturing defect detection using VGG16 convolutional neural networks. Wireless Communications and Mobile Computing, 2022, Article ID 1070405. https://doi.org/10.1155/2022/1070405

[2] Biradar, M. S., Shiparamatti, B. G., & Patil, P. M. (2021). Fabric defect detection using deep convolutional neural network. Optical Memory and Neural Networks, 30(3), 250–256. https://doi.org/10.3103/S1060992X21030024

[3] Block, S. B., da Silva, R. D., Dorini, L. B., & Minetto, R. (2021). Inspection of imprint defects in stamped metal surfaces using deep learning and tracking. IEEE Transactions on Industrial Electronics, 68(5), 4498–4507. https://doi.org/10.1109/TIE.2020.2993526

[4] Božič, J., Tabernik, D., & Skočaj, D. (2021). Mixed supervision for surface-defect detection: From weakly to fully supervised learning. Computers in Industry, 129, 103459. https://doi.org/10.1016/j.compind.2021.103459

[5] Cumbajin, E., Rodrigues, N., Costa, P., Miragaia, R., Frazão, L., Costa, N., … & Pereira, A. (2023). A systematic review on deep learning with CNNs applied to surface defect detection. Journal of Imaging, 9(10), 193. https://doi.org/10.3390/jimaging9100193

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

[7] Kahraman, Y., & Durmuşoğlu, A. (2023). Deep learning-based fabric defect detection: A review. Textile Research Journal, 93(12), 1485–1503. https://doi.org/10.1177/00405175221130773

[8] Omarov, B., Suliman, A., Tsoy, A. Parallel backpropagation neural network training for face recognition. Far East Journal of Electronics and Communications. Volume 16, Issue 4, December 2016, Pages 801-808. (2016).

[9] Lin, H. I., & Wibowo, F. S. (2021). Image data assessment approach for deep learning-based metal surface defect-detection systems. IEEE Access, 9, 47621–47638. https://doi.org/10.1109/ACCESS.2021.3068478

[10] Altayeva, A., Omarov, B., Suleimenov, Z., & Im Cho, Y. (2017, June). Application of multi-agent control systems in energy-efficient intelligent building. In 2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS) (pp. 1-5). IEEE.

[11] Baiganova, A., Toxanova, S., Yerekesheva, M., Nauryzova, N., Zhumagalieva, Z., & Tulendi, A. (2024). Hybrid Convolutional Recurrent Neural Network for Cyberbullying Detection on Textual Data. International Journal of Advanced Computer Science & Applications, 15(5).

[12] Omarov, B., Omarov, B., Shekerbekova, S., Gusmanova, F., Oshanova, N., Sarbasova, A., ... & Sultan, D. (2019). Applying face recognition in video surveillance security systems. In Software Technology: Methods and Tools: 51st International Conference, TOOLS 2019, Innopolis, Russia, October 15–17, 2019, Proceedings 51 (pp. 271-280). Springer International Publishing.

[13] Markatos, N. G., & Mousavi, A. (2023). Manufacturing quality assessment in the Industry 4.0 era: A review. Total Quality Management & Business Excellence. Advance online publication. https://doi.org/10.1080/14783363.2023.2194524

[14] Patil, D. B., Nigam, A., Mohapatra, S., & Nikam, S. (2023). A deep learning approach to classify and detect defects in the components manufactured by laser directed energy deposition process. Machines, 11(9), 854. https://doi.org/10.3390/machines11090854

[15] Pathak, K. A., Kafle, P., & Vikram, A. (2025). Deep learning-based defect detection in film-coated tablets using a convolutional neural network. International Journal of Pharmaceutics, 671, 125220. https://doi.org/10.1016/j.ijpharm.2025.125220

[16] Prunella, M., Scardigno, R. M., Buongiorno, D., Brunetti, A., Longo, N., Carli, R., … & Bevilacqua, V. (2023). Deep learning for automatic vision-based recognition of industrial surface defects: A survey. IEEE Access, 11, 43370–43423. https://doi.org/10.1109/ACCESS.2023.3271748

[17] Profili, A., Magherini, R., Servi, M., Spezia, F., Gemmiti, D., & Volpe, Y. (2024). Machine vision system for automatic defect detection of ultrasound probes. The International Journal of Advanced Manufacturing Technology. Advance online publication. https://doi.org/10.1007/s00170-024-14701-6

[18] Saberironaghi, A., Ren, J., & El-Gindy, M. (2023). Defect detection methods for industrial products using deep learning techniques: A review. Algorithms, 16(2), 95. https://doi.org/10.3390/a16020095

[19] Shahrabadi, S., Castilla, Y., Guevara, M., Magalhães, L. G., Gonzalez, D., & Adão, T. (2022). Defect detection in the textile industry using image-based machine learning methods: A brief review. Journal of Physics: Conference Series, 2224(1), 012010. https://doi.org/10.1088/1742-6596/2224/1/012010

[20] Al Noman, M. A., Zhai, L., Almukhtar, F. H., Rahaman, M. F., Omarov, B., Ray, S., ... & Wang, C. (2023). A computer vision-based lane detection technique using gradient threshold and hue-lightness-saturation value for an autonomous vehicle. International Journal of Electrical and Computer Engineering, 13(1), 347.

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

[22] Wan, P. K., & Leirmo, T. L. (2023). Human-centric zero-defect manufacturing: State-of-the-art review, perspectives, and challenges. Computers in Industry, 144, 103792. https://doi.org/10.1016/j.compind.2022.103792

[23] Albanese, A., Nardello, M., Fiacco, G., & Brunelli, D. (2023). Tiny machine learning for high accuracy product quality inspection. IEEE Sensors Journal, 23(2), 1575–1583. https://doi.org/10.1109/JSEN.2022.3140084

[24] Li, D., Hua, S., Li, Z., Gong, X., & Wang, J. (2022). Automatic vision-based online inspection system for broken-filament of carbon fiber with multiscale feature learning. IEEE Transactions on Instrumentation and Measurement, 71, 1–12. https://doi.org/10.1109/TIM.2022.3154818

[25] Kumaresan, S., Aultrin, K. S. J., Kumar, S. S., & Dev Anand, M. (2023). Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning. International Journal on Interactive Design and Manufacturing, 17(4), 2999–3010. https://doi.org/10.1007/s12008-023-01327-3

[26] Pranoto, K. A., Caesarendra, W., Tjahjowidodo, T., & Lim, G. H. (2023). Burrs and sharp edge detection of metal workpiece using CNN image

classification method for intelligent manufacturing applications. In 2023 IEEE 21st International Conference on Industrial Informatics (INDIN) (pp. 1–7). https://doi.org/10.1109/INDIN55582.2023.10196164

[27] Smagulova, D., Samaitis, V., & Jasiuniene, E. (2024). Convolutional neural network for interface defect detection in adhesively bonded dissimilar structures. Applied Sciences, 14(22), 10351. https://doi.org/10.3390/app142210351

[28] Li, Y., Gao, P., Luo, Y., Luo, X., Xu, C., Chen, J., … & Xu, W. (2024). Automatic detection and classification of natural weld defects using alternating magneto-optical imaging and ResNet50. Sensors, 24(23), 7649. https://doi.org/10.3390/s24237649

[29] Kumar, N., & Kumar, D. (2022). Deep learning methods for object detection in smart manufacturing: A comprehensive survey. Journal of Manufacturing Systems, 65, 424–445. https://doi.org/10.1016/j.jmsy.2022.02.008

[30] Omarov, B., Batyrbekov, A., Suliman, A., Omarov, B., Sabdenbekov, Y., & Aknazarov, S. (2020, November). Electronic stethoscope for detecting heart abnormalities in athletes. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-5). IEEE.

[31] Omarov, B., Batyrbekov, A., Dalbekova, K., Abdulkarimova, G., Berkimbaeva, S., Kenzhegulova, S., ... & Omarov, B. (2021). Electronic stethoscope for heartbeat abnormality detection. In Smart Computing and Communication: 5th International Conference, SmartCom 2020, Paris, France, December 29–31, 2020, Proceedings 5 (pp. 248-258). Springer International Publishing.

[32] Mattern, A., Gerdes, H., Grunert, D., & Schmitt, R. H. (2025). A comparison of transformer and CNN-based object detection models for surface defects on Li-Ion battery electrodes. Journal of Energy Storage, 105, 114378. https://doi.org/10.1016/j.est.2025.114378

[33] Chen, X., Zhang, X., Shi, Y., & Pang, J. (2025). HCT-Det: A high-accuracy end-to-end model for steel defect detection based on hierarchical CNN–Transformer features. Sensors, 25(5), 1333. https://doi.org/10.3390/s25051333

[34] Raj, G. D., & Prabadevi, B. (2025). Enhancing surface detection: A comprehensive analysis of various YOLO models. Heliyon, 11(3), e42433. https://doi.org/10.1016/j.heliyon.2025.e42433

[35] Szőlősi, J., Magyar, P., Bán, A., et al. (2024). Welding defect detection with image processing on a custom small dataset: A comparative study. IET Collaborative Intelligent Manufacturing. https://doi.org/10.1049/cim2.12093

[36] Kumaresan, S., Jai Aultrin, K. S., Kumar, S. S., & Dev Anand, M. (2023). Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning. International Journal on Interactive Design and Manufacturing, 17, 2999–3010. https://doi.org/10.1007/s12008-023-01327-3

[37] Li, N., Wang, Z., Zhao, R., Yang, K., & Ouyang, R. (2025). YOLO-PDC: An improved YOLOv7 model for aluminum surface defect detection. Journal of Real-Time Image Processing, 22, 86

[38] Aslam, Y., Santhi, N., Ramasamy, N., & Ramar, K. (2021). Localization and segmentation of metal cracks using deep learning. Journal of Ambient Intelligence and Humanized Computing, 12(5), 4205–4213. https://doi.org/10.1007/s12652-020-02580-7

[39] Shafi, I., Mazhar, M. F., Fatima, A., Alvarez, R. M., Miró, Y., Martínez Espinosa, J. C., & Ashraf, I. (2023). Deep learning-based real time defect detection for optimization of aircraft manufacturing and control performance. Drones, 7(1), 31. https://doi.org/10.3390/drones7010031

[40] Kazmi, S., O'Shea, D., & Walsh, J. (2023). A deep learning-based framework for visual inspection of plastic bottles in an Industry 4.0 context. IEEE Access, 11, 125529–125542. https://doi.org/10.1109/ACCESS.2023.3307958

[41] Hussain, M., Chen, T., Titrenko, S., Su, P., & Mahmud, M. (2022). A gradient-guided architecture coupled with filter-fused representations for micro-crack detection in photovoltaic cell surfaces. IEEE Access, 10, 58950–58964. https://doi.org/10.1109/ACCESS.2022.3178675

[42] Zahid, A., Hussain, M., Hill, R., & Al-Aqrabi, H. (2023, May). Lightweight convolutional network for automated photovoltaic defect detection. In 2023 9th International Conference on Information Technology Trends (ITT) (pp. 133–138). https://doi.org/10.1109/ITT57172.2023.10187264

[43] Xu, Y., Zhang, K., & Wang, L. (2021). Metal surface defect detection using modified YOLO. Algorithms, 14(9), 257. https://doi.org/10.3390/a14090257

[44] Baikuvekov, M., Tursynova, A., & Yespayev, G. (2024, May). A Deep Learning for Cardiovascular Diseases Detection on Wearable Devices Data. In 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST) (pp. 272-277). IEEE.

[45] Tileubay, S., Yerekeshova, M., Baiganova, A., Janyssova, D., Omarov, N., Omarov, B., & Baiekeyeva, Z. (2024). Development of Deep Learning Enabled Augmented Reality Framework for Monitoring the Physical Quality Training of Future Trainers-Teachers. International Journal of Advanced Computer Science & Applications, 15(3).

[46] Tang, Y., Chen, X., & Yang, J. (2023). Surface defect detection of bearing rings based on an improved YOLOv5. Machines, 11(4), 469. https://doi.org/10.3390/machines11040469

[47] Tursynova, A., & Kaldarova, B. (2024). Diagnóstico precoz de accidentes cerebrovasculares en atletas de halterofilia en tiempo real utilizando sensores no invasivos de última generación. Retos, 61, 1321–1332. https://doi.org/10.47197/retos.v61.110267

[48] Li, P., Wen, S., Zhao, D., Huang, X., Liu, Z., & Guo, L. (2023). Adaptive detection of multi-scale casting defects via a global dynamic transformer. Computers in Industry, 146, 103870. https://doi.org/10.1016/j.compind.2023.103870

[49] Baek, D., Moon, H. S., & Park, S. H. (2021). Deep learning-based defects detection in keyhole TIG welding processes. Journal of Welding and Joining, 39(6), 565–574. https://doi.org/10.5781/JWJ.2021.39.6.8

[50] Park, S.-H., Lee, K.-H., Park, J.-S., & Shin, Y.-S. (2022). Deep learning-based defect detection for sustainable smart manufacturing. Sustainability, 14(5), 2697. https://doi.org/10.3390/su14052697

[51] Baek, D., Kim, J., & Moon, H. S. (2022). Deep learning-based defect detection for hot-rolled strip steel surfaces. Journal of Physics: Conference Series, 2246(1), 012080. https://doi.org/10.1088/1742-6596/2246/1/012080

[52] Zhang, Z., Li, X., & Wang, Y. (2023). An ensemble-based deep learning model for welding defect detection in submerged arc welds. Journal of Manufacturing Processes, 92, 83–93. https://doi.org/10.1016/j.jmapro.2023.07.033