Society of Production Engineering

ICPES 2025
40th INTERNATIONAL CONFERENCE ON PRODUCTION ENGINEERING - SERBIA 2025

DOI: 10.46793/ICPES25.189GM

University of Nis Faculty of Mechanical Engineering

Nis, Serbia, 18 - 19th September 2025

DEEP LEARNING TECHNIQUES FOR DEFECT DETECTION IN AUTOMATED QUALITY CONTROL SYSTEMS

A. Gandhi MANIKANDAN[1*], V. SANDHIYA[1], P. HARINE[1], R. AJAY DHISONE[1], M. SATHYAPRAKASH[1].

…; Orcid: 0009-0004-7979-2721**; Orcid:** 0009-0006-7509-9254; **…..; …..**

[1]SRM TRP Engineering College, Trichy, Tamilnadu, India - 621105

*Corresponding author: nagadeepan.a@gmail.com

Abstract: *Automated quality control (AQC) systems have transformed modern manufacturing by enabling high-speed, consistent, and cost-efficient inspection of products. However, traditional image processing techniques often fail to detect subtle, non-conforming defects due to limitations in adaptability and generalization. This paper explores the integration of deep learning (DL) techniques—particularly convolutional neural networks (CNNs), autoencoders, and vision transformers—for enhancing defect detection accuracy in AQC environments. A comprehensive methodology is presented where training datasets are augmented using synthetic defect generation, followed by supervised and unsupervised learning approaches for feature extraction and classification. The research evaluates various model architectures using metrics such as accuracy, precision, recall, and inference time on datasets collected from electronic component inspection and surface defect analysis in metal casting. The results demonstrate that DL models significantly outperform conventional rule-based systems, particularly in detecting micro-defects and anomalies in complex textures. Moreover, the incorporation of transfer learning and model pruning techniques further reduces computational overhead, making the deployment feasible in real-time production lines. This study concludes with an outline of implementation challenges, including data imbalance, hardware constraints, and model interpretability, and proposes potential directions for future research in intelligent adaptive inspection systems. The findings aim to contribute toward the development of robust, scalable, and intelligent quality control frameworks aligned with Industry 4.0 principles.*

Keywords: *deep learning, defect detection, automated quality control, convolutional neural networks, computer vision, real-time inspection, Industry 4.0*

1. INTRODUCTION

The rise of Industry 4.0 has necessitated a paradigm shift in manufacturing processes, with a strong emphasis on automation, interconnectedness, and intelligent systems. Automated Quality Control (AQC) is a critical component of this transformation, moving away from manual and error-prone inspection methods to fast, reliable, and objective systems. Traditional AQC systems often rely on classical computer vision techniques, such as thresholding, edge detection, and feature matching [1,2]. While effective for simple, well-defined defects, these methods struggle with complex, textured surfaces and subtle, non-conforming anomalies. They lack the adaptability and generalization capabilities required to handle the wide variety of defects that can occur in real-world production environments.

This paper addresses the limitations of conventional AQC systems by proposing the integration of advanced deep learning (DL) techniques for defect detection. DL models, particularly those based on neural networks, have demonstrated state-of-the-art performance in complex image recognition tasks, making them ideal candidates for enhancing the accuracy and robustness of AQC. This study focuses on three key DL architectures: Convolutional Neural Networks (CNNs), Autoencoders, and Vision Transformers (ViTs). We present a comprehensive methodology for dataset preparation, model training, and performance evaluation, and discuss the results of their application to real-world manufacturing datasets. The ultimate goal is to contribute to the development of intelligent quality control frameworks that are scalable and aligned with the principles of Industry 4.0.

2. METHODOLOGY

Our methodology is structured to provide a robust and systematic approach to deep learning-based defect detection. The process begins with dataset acquisition, followed by data augmentation, model selection and training, and finally, performance evaluation.

2.1 Dataset Preparation

We utilized two distinct datasets to evaluate the models: one consisting of electronic components and another comprising images of metal castings. The electronic component dataset includes various micro-defects such as solder bridges, missing components, and cracks. The metal casting dataset contains surface anomalies like pores, cracks, and scratches. To address the typical data imbalance problem where defect samples are rare, we employed data augmentation techniques. These included standard image transformations (rotation, scaling, flipping) as well as more advanced methods like synthetic defect generation using Generative Adversarial Networks (GANs) or image manipulation techniques to create realistic defect examples.

Model Architectures and Training

Three primary deep learning architectures were chosen for this study:

1. Convolutional Neural Networks (CNNs): CNNs, such as VGGNet, ResNet, and YOLO, are well-suited for feature extraction from images. We trained several CNN models in a supervised manner, where each image was labeled as either “defective” or “non-defective,” or with specific defect types.

1. Autoencoders:
   a) For unsupervised learning, we used a standard autoencoder and a variational autoencoder (VAE) [3]. These models are trained to reconstruct normal, non-defective images. Defects are identified by a high reconstruction error, as the model struggles to reproduce features it has not been trained on. This approach is particularly useful in scenarios with very few defect samples.
2. Vision Transformers (ViTs):
   a) ViTs represent a more recent approach, adapted from natural language processing. They process images as sequences of patches and use a self-attention mechanism to

learn global dependencies [4]. We fine-tuned a pre-trained ViT model for the defect classification task, which leverages the benefits of transfer learning.

Model Evaluation :

To rigorously assess the effectiveness of each deep learning architecture deployed in this study, a comprehensive evaluation framework was employed. The models were benchmarked using a suite of standard classification metrics—accuracy, precision, recall, and F1-score—which collectively provide a multidimensional view of performance across both balanced and imbalanced datasets.

Accuracy reflects the overall correctness of the model's predictions, offering a general measure of performance.

Precision quantifies the proportion of true positive predictions among all positive predictions, which is critical in minimizing false alarms in defect detection.

Recall measures the model's ability to identify all actual defect instances, ensuring that no defective item escapes detection.

F1-score, the harmonic mean of precision and recall, balances the trade-off between false positives and false negatives, making it particularly valuable in quality control contexts where both types of errors carry operational consequences.

In addition to these classification metrics, inference time—defined as the time required for the model to process and classify a single image—was measured to evaluate the feasibility of real-time deployment. This metric is especially pertinent in high-throughput manufacturing environments where latency can directly impact production efficiency and decision-making speed.

All performance metrics were computed on a separate, unseen test dataset that was not used during training or validation. This approach ensures that the evaluation reflects the model's generalization capability, i.e., its ability to perform reliably on new, previously unencountered data. Such validation is essential for industrial applications, where models must maintain robustness across varying lighting conditions, surface textures, and defect types.

3. RESULT AND DISCUSSION :

The empirical findings of this study affirm the superior performance of deep learning models over traditional image processing techniques in the domain of automated quality control (AQC). By leveraging hierarchical feature extraction and nonlinear representation learning, deep neural networks—particularly Convolutional Neural Networks (CNNs), autoencoders, and Vision Transformers (ViTs)—demonstrated a remarkable ability to detect subtle and complex defect patterns that conventional algorithms often fail to capture.

3.1 Electronic Component Inspection

In the task of inspecting electronic components, the ResNet-50 architecture emerged as a high-performing model, achieving an outstanding accuracy of 98.5% and precision of 97.2%. These metrics represent a substantial leap over traditional methods such as Canny edge detection, which yielded only 85.1% accuracy, highlighting the limitations of rule-based feature extraction in handling nuanced defect morphologies.

The Vision Transformer (ViT), enhanced through transfer learning, exhibited exceptional proficiency in identifying texture-based anomalies. Its global attention mechanism enabled the model to capture long-range dependencies and contextual cues, making it particularly effective in scenarios where defect patterns are spatially dispersed or irregular.

3.2 Metal Casting Surface Analysis

For surface defect detection in metal casting applications, an unsupervised autoencoder model was employed. Trained exclusively on defect-free samples, the autoencoder learned a compact representation of normal surface textures. During inference, deviations from this learned manifold were flagged as anomalies. This approach yielded a recall rate of 95.8%, underscoring its efficacy in identifying rare or previously unseen defects without requiring extensive labeled datasets.

Such anomaly detection frameworks are especially valuable in industrial contexts where defect occurrences are infrequent, and the cost or feasibility of collecting large-scale annotated data is prohibitive.

4. IMPLEMENTATION CHALLENGES

Despite the promising performance of deep learning-based AQC systems, several practical and infrastructural challenges must be addressed to facilitate their widespread adoption in manufacturing environments.

4.1 Data Imbalance and Scarcity

One of the most pressing issues is the imbalance and scarcity of real-world defect data. In many industrial settings, defective samples constitute a small fraction of the overall production output, leading to skewed datasets that hinder model generalization. While synthetic data generation techniques—such as image augmentation, GAN-based synthesis, and simulation-driven rendering—offer partial remedies, they often fall short in replicating the full spectrum of defect variability encountered in practice.

Future research should explore domain adaptation, few-shot learning, and active learning strategies to enhance model robustness under data-constrained conditions.

4.2 Hardware and Computational Constraints

The deployment of deep learning models for real-time quality control necessitates high-throughput inference capabilities, which in turn require specialized hardware such as Graphics Processing Units (GPUs) or edge AI accelerators. For small and medium-sized enterprises (SMEs), the financial and infrastructural investment needed to support such hardware can be prohibitive.

Efforts toward model compression, quantization, and efficient architecture design (e.g., MobileNet, EfficientNet) are essential to enable cost-effective deployment without compromising performance. Additionally, cloud-based inference and federated learning frameworks may offer scalable alternatives for resource-limited settings.

5. CONCLUSION

This study has demonstrated the substantial promise of deep learning techniques—namely Convolutional Neural Networks (CNNs), autoencoders, and Vision Transformers—in reshaping automated quality control systems across diverse industrial applications. These models not only surpass traditional image processing methods in terms of defect detection accuracy and robustness, but also offer adaptive learning capabilities that are critical for dynamic manufacturing environments.

By incorporating strategies such as transfer learning and model pruning, the proposed frameworks achieve both high performance and computational efficiency, making them suitable for real-time deployment. The success of unsupervised approaches, particularly in scenarios with limited defect data, further underscores the versatility of deep learning in addressing practical constraints faced by manufacturers.

Looking forward, the evolution of intelligent quality control systems will depend on several key research directions: the development of advanced data augmentation techniques to mitigate data imbalance; the design of hybrid architectures that synergize the strengths of multiple models; and the integration of explainable AI (XAI) to enhance transparency, trust, and regulatory compliance.

ACKNOWLEDGEMENT

This research was supported by the SRM TRP Engineering College. We extend our gratitude to our colleagues for their valuable feedback and assistance.

REFERENCES

[1] H. Czichos, D. Klaffke, E. Santner, M. Woydt: Advances in tribology: the materials point of view, Wear, Vol. 190, No. 2, pp. 155-161, 1995.

[2] D. P. Kingma and M. Welling: Auto-Encoding Variational Bayes, in: International Conference on Learning Representations, 2014.

[3] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, in: International Conference on Learning Representations, 2021.