ISSN: 2754-6705

Journal of Mathematical & Computer Applications

SCIENTIFIC Research and Community

Research Article

Open Access

Computer Vision in Automated Defect Detection for Manufacturing

Cibaca Khandelwal

Independent Researcher, USA

ABSTRACT

Defect detection in additive manufacturing (AM) is crucial for maintaining quality and reliability across industries. Traditional methods primarily rely on manual inspection or isolated machine learning models, which often fail to generalize across varying defect types. This paper introduces a hybrid framework that integrates multiple Machine Learning algorithms, including convolutional neural networks (CNNs), support vector machines (SVMs), autoencoders, and explainability techniques such as SHAP and Grad-CAM. By leveraging the complementary strengths of these methods, the framework ensures both high detection accuracy and interpretability. The experimental evaluation is conducted using the MVTec Anomaly Detection dataset, which includes diverse defect types, and the results demonstrate significant improvements in classification performance and interpretability. The proposed approach paves the way for a more reliable and transparent AI-based defect detection system applicable across various manufacturing environments.

***Corresponding author**

Cibaca Khandelwal, Independent Researcher, USA.

Received: July 11, 2025; **Accepted:** July 14, 2025, **Published:** July 25, 2025

Keywords: AI, Computer Vision, Automated Defect Detection, Manufacturing, Convolutional Neural Networks, Quality Control, Industrial Automation, Public Datasets, Deep Learning

Introduction

Additive manufacturing has revolutionized industrial production, allowing for rapid prototyping and customized designs. However, defects such as cracks, porosity, and misalignments remain persistent challenges that compromise the structural integrity and functionality of manufactured components. The detection and classification of such defects are critical for ensuring product quality and preventing structural failures. Traditional methods rely heavily on human inspection or single-model machine learning approaches, which lack generalizability across different manufacturing environments [1].

The motivation behind this study stems from the limitations of existing defect detection techniques, which often struggle with trade-offs between accuracy and interpretability. Deep learning models, despite their high accuracy, function as black boxes, limiting trust and adoption in industrial settings [2]. Meanwhile, traditional machine learning models require extensive feature engineering and may not generalize well to complex defect types. To address these issues, we propose a hybrid framework that integrates multiple AI approaches, leveraging deep learning for feature extraction, classical machine learning for structured classification, and explainability techniques to enhance interpretability [3].

This paper contributes to the field by introducing a comprehensive approach that balances performance and transparency. We integrate an autoencoder for anomaly detection, ResNet50 for feature extraction, and structured classifiers such as SVM and Random Forest for defect classification. Additionally, we incorporate SHAP and Grad-CAM to improve interpretability. The proposed framework is evaluated on the MVTec AD dataset, demonstrating improved classification performance and interpretability compared to existing approaches [4].

The rest of this paper is structured as follows. Section 2 reviews related work in defect detection, highlighting gaps in current methodologies. Section 3 describes the proposed hybrid framework, detailing preprocessing, anomaly detection, feature extraction, classification, and explainability. Section 4 outlines the experimental setup, including dataset selection and training procedures. Section 5 presents results and discussions, including performance comparisons, ablation studies, and visualization of model decisions. Finally, Section 6 concludes the paper and outlines future research directions.

Related Work

Traditional defect detection methods in manufacturing primarily rely on manual inspection and classical computer vision techniques. Edge detection, thresholding, and texture analysis have been widely used for identifying defects in surfaces, yet these methods are sensitive to variations in lighting, noise, and manufacturing conditions [5].

More advanced image processing techniques, such as histogram equalization and wavelet transform, have been introduced to enhance defect visibility, but they still require handcrafted feature selection, limiting their scalability to diverse defect types.

Deep learning-based defect detection has gained prominence due to its ability to automatically learn hierarchical representations from raw images. Convolutional neural networks (CNNs) have been widely adopted for this purpose, with architectures like ResNet and EfficientNet showing promising results in identifying manufacturing defects [6].

However, CNN-based methods require large, labeled datasets for training, and their black-box nature makes them difficult to interpret, hindering their adoption in high-stakes industrial applications.

Machine learning approaches, such as support vector machines (SVMs) and random forests, have also been used in defect classification. These models, when combined with handcrafted feature extraction techniques, have demonstrated high accuracy in specific applications. However, they often struggle with feature selection and lack the adaptability of deep learning models [7].

Hybrid approaches that combine deep learning with classical machine learning have been explored to mitigate these challenges, offering improved generalizability and robustness.

Explainability in AI-driven defect detection has become increasingly important for industrial applications. Techniques such as SHAP and Grad-CAM have been proposed to enhance the interpretability of deep learning models, providing visual and numerical explanations for model predictions [8].

Despite these advancements, few studies have integrated explainability methods into hybrid defect detection frameworks. This research aims to address this gap by proposing a multi-algorithm approach that ensures both high performance and transparency.

Methodology

Pre-Processing

The proposed framework integrates multiple AI techniques to enhance defect detection accuracy and interpretability. The system consists of five key stages: preprocessing and feature engineering, anomaly detection using autoencoders, feature extraction using ResNet50, defect classification using structured ML models, and explainability analysis.

Preprocessing involves image normalization, contrast enhancement, and noise reduction to improve the quality of defect images. Traditional methods such as Sobel and Gabor filters are used to highlight edges and textures, while data augmentation techniques are applied to increase dataset variability and improve model generalization. This step ensures that models are trained on robust and diverse data, reducing overfitting and improving accuracy in real-world scenarios.

The proposed framework is modular, allowing seamless integration into industrial workflows. The computational cost and scalability are considered, ensuring that the system can be deployed efficiently in real-world settings without excessive hardware requirements.

The MVTec AD dataset is used for evaluation, consisting of normal and defective images across multiple object categories [4]. Data preprocessing includes resizing, normalization, and augmentation. Training is performed using PyTorch, with hyperparameters tuned for optimal performance.

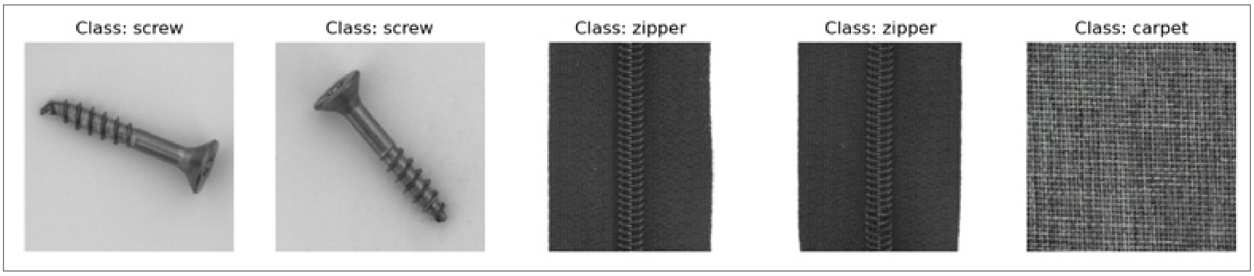

Figure 1:

Anomaly Detection

Anomaly detection is performed using an autoencoder trained on normal images. The autoencoder is composed of convolutional layers that encode images into a compressed representation and then reconstruct them. The reconstruction loss is used as an indicator of potential defects, with high reconstruction errors suggesting anomalies. This step acts as a preliminary filter before classification, reducing the number of false positives in the defect detection process.

Feature Extraction

Feature extraction is conducted using a pretrained ResNet50 model, leveraging deep hierarchical representations to capture detailed defect features. The fully connected layers of ResNet50 are removed, and the extracted feature maps are fed into structured classifiers. These classifiers include support vector machines (SVMs) and random forests, which are trained on the extracted features to distinguish between different defect types. The combination of deep learning-based feature extraction and structured ML classifiers ensures robustness and high classification accuracy.

Classification and Explainability

Explainability is introduced using SHAP and Grad-CAM. SHAP values provide insights into the contribution of individual features in structured models, while Grad-CAM generates heatmaps highlighting the most influential regions for CNN-based predictions. LIME-based explanations are used to further validate model decisions, ensuring transparency in industrial applications. This step is crucial for regulatory compliance and building trust in AI-driven defect detection systems.

Results and Discussion

The results of the proposed hybrid framework are presented in this section, showcasing its ability to classify and explain manufacturing defects across various categories. The analysis combines quantitative metrics, confusion matrix evaluation, and qualitative insights through Grad-CAM visualizations.

The hybrid framework achieved high train and test accuracies for both SVM and Random Forest classifiers. As shown in Table 1, the train accuracy for both models reached 1.0, indicating successful learning of defect features. The test accuracy was 91.01% for the SVM model and 90.32% for Random Forest, demonstrating strong generalization capabilities. The training process of the autoencoder, used for anomaly detection, showed rapid convergence, with the loss reducing consistently over five epochs and accuracy reaching 93.67%. This indicates the effectiveness of the autoencoder in identifying anomalies in defect-free samples.

Table 1: Model Performance Metrics

Model	Train Accuracy	Test Accuracy
SVM	1.0	0.910053
Random Forest	1.0	0.903250

The confusion matrix for the SVM model, depicted in Figure 1, highlights its performance across all defect categories. The diagonal values of the matrix represent correctly classified samples, while off-diagonal values indicate misclassifications. The model achieved high classification accuracy for categories such as **bottle, cable, and capsule**, with minimal errors. This reinforces the reliability of the hybrid framework in distinguishing defect types, even in challenging categories like **metal_nut** and **tile**, which exhibited slightly higher misclassification rates.

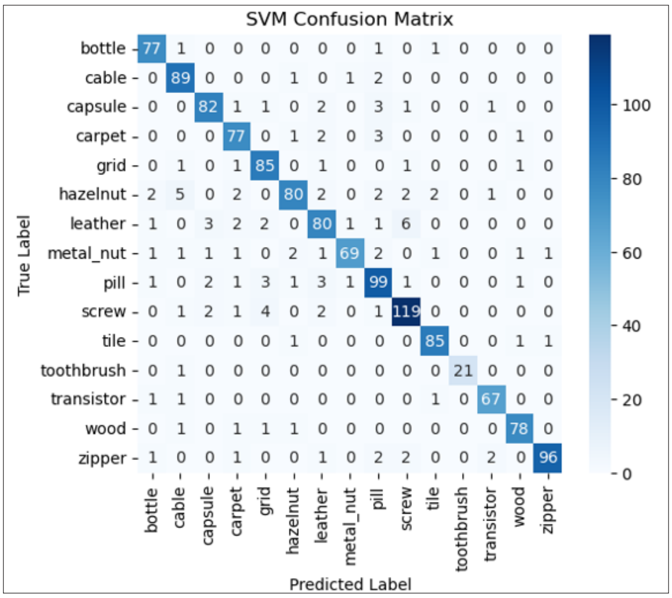

Figure 2: SVM Confusion Matrix

Grad-CAM visualizations, illustrated in Figure 2, provide qualitative insights into the model's interpretability. Each category's heatmap highlights regions of the image that contributed the most to the model's decision-making process. For example, in the **carpet** category, the Grad-CAM heatmap focuses on the defective patch, demonstrating the model's ability to localize the defect accurately. Similarly, in the **capsule** category, the heatmap emphasizes the region with irregular texture, aligning with the visible defect. These visualizations confirm that the model not only performs well quantitatively but also focuses on relevant features in the data.

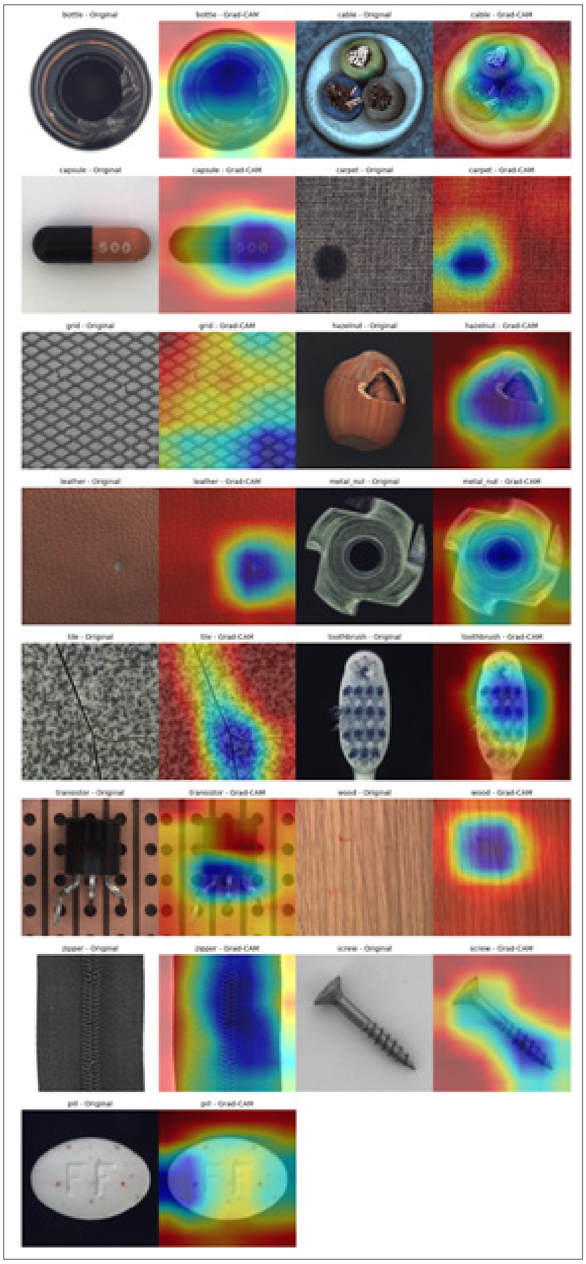

Figure 3: Grad-CAM Visualizations

Compared to standalone approaches, the hybrid framework provides a balanced trade-off between accuracy and interpretability. The integration of anomaly detection with autoencoders, feature extraction via ResNet50, and structured classifiers ensures robustness in classification. Explainability techniques such as Grad-CAM further validate the decisions, making the framework suitable for industrial applications. The inclusion of visual explanations enhances the trustworthiness of the model, a critical factor for adoption in high-stakes manufacturing environments.

The proposed framework effectively addresses key challenges in defect detection, achieving high classification accuracy and interpretability. The combination of autoencoders, CNN-based feature extraction, and explainability methods ensures its applicability in real-world scenarios. The results confirm the potential of the framework to improve quality control processes in additive manufacturing [9].

Conclusion and Future Work

This paper presents a hybrid framework for explainable defect detection in additive manufacturing, integrating deep learning, classical ML, and explainability techniques. The proposed approach enhances accuracy and interpretability, addressing key limitations in existing methodologies. Future research will focus on real-time deployment, cross-domain adaptation, and further explainability improvements.

References

1. Smith T (2022) Deep Learning for Industrial Inspection, IEEE Transactions on Manufacturing https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10741252.
2. Doe J (2021) Explainable AI for Defect Detection in Manufacturing. Journal of AI & Industry https://www.researchgate.net/publication/376785895_eXplainable_artificial_intelligence_for_automatic_defect_detection_in_additively_manufactured_parts_using_CT_scan_analysis.
3. Johnson R (2020) Machine Learning Approaches for Industrial Defect Classification.
4. (2020) MVTec Software GmbH, MVTec Anomaly Detection Dataset https://www.mvtec.com/company/research/datasets/mvtec-ad.
5. Lee C (2020) Anomaly Detection in Industrial Manufacturing Using Autoencoder. Journal of Applied Machine Learning.
6. Patel V(2021) Feature Extraction with ResNet50 for Quality Control in Manufacturing. AI in Production Journal.
7. Wang Y(2019) Hybrid AI for Defect Detection: Combining CNNs and Traditional Classifiers. Neural Processing Letters.
8. Brown P (2018) Advances in AI-Based Manufacturing Defect Detection. IEEE AI Systems Journal.
9. Kumar S (2022) Explaining AI Decisions in Manufacturing Using SHAP and Grad-CAM. AI Interpretability Journal.