

Research Paper

CASTGUARD VISION: SMART DEFECT RECOGNITION IN MANUFACTURING INSPECTION SYSTEMS

BOYINA GOPI RAJU¹, PULI YOGESWARI²

¹ HOD & Associate Professor, Department of MCA, Chirala Engineering College, Chirala, Andhra Pradesh, India-523157.

Email: rajuboyina001@gmail.com

² MCA Student, Department of MCA, Chirala Engineering College, Chirala, Andhra Pradesh, India-523157.

Email: puliyogeswari@gmail.com

ABSTRACT

The rapid growth of industrial automation has increased the need for reliable and efficient quality inspection systems, particularly in manufacturing sectors such as casting production. Traditional inspection methods rely heavily on manual observation, making them time-consuming, inconsistent, and prone to errors caused by human fatigue and subjective judgment. With the increasing availability of image-based industrial data, machine learning techniques provide a promising solution for automating defect detection and improving inspection accuracy. This study focuses on the identification and classification of defects in casting product images. Conventional inspection systems often struggle to maintain consistency and scalability when processing large volumes of images. Existing approaches primarily depend on human expertise or rule-based image processing methods, which require handcrafted features and lack adaptability to varying industrial conditions. As a result, these systems often fail to detect subtle defects accurately and cannot efficiently handle large datasets. To address these limitations, a machine learning-based image classification system is proposed using Multinomial Naive Bayes (MNB), Decision Tree Classifier (DTC), and Random Forest Classifier (RFC). The system performs image preprocessing through resizing, normalization, and feature extraction by converting images into flattened feature vectors. These features are used to train and evaluate the models, with performance measured using accuracy, precision, recall, and F1-score. The proposed system enhances defect detection by providing automated, accurate, and consistent classification. The use of Random Forest, an ensemble learning technique, improves prediction performance while reducing overfitting. Additionally, a graphical user interface (GUI) enables easy user interaction, making the system suitable for practical industrial applications. Overall, the proposed approach offers a scalable, reliable, and efficient solution for casting defect detection, contributing to improved quality control and operational efficiency.

Keywords: Inspection Systems, Machine Learning, Random Forest Classifier, Graphical User Interface, Image Processing.

1. INTRODUCTION

Casting defect detection is an essential problem in the machinery industry. High

quality castings are very important for automobiles, engineering equipment and other products. It is necessary to find and deal with these defective castings in time as shown in

Fig.1. Otherwise, it will harm the quality of downstream industrial products [1]. With the development of computer vision technology, defects detection methods based on casting image processing have been proposed and applied in recent years. These progresses are made upon various excellent network structures.[2]. In the past decade, in order to improve the classification accuracy, the models are designed to be more and more complex, which leads to an increasingly large number of parameters and an exponential increase in the computational burdens, especially for large images [3]. The image size of the castings is also large, but usually the defects are only located in a small number of the pixels, which means large numbers of the pixels in the image are useless for the defect's detection task, but they still consume equivalent computational resources.



Fig. 1: Casting product quality inspection process.

Image classification based on reinforcement learning has been a research hot spot in recent years. In the work of Ref. [4], the policy gradient method, which is a very basic reinforcement learning algorithm, is used to optimize the convolutional neural network model for image classification. On the contrary, reinforcement learning algorithms can also be used as an effective tool to find the key region in the image. In Ref. [5], Deep Q-learning algorithm is used to select a vital area of the image in a vehicle classification task. In Ref, a multi-agent algorithm PixelRL, which is based on the A3C [6,7,8] algorithm, is

proposed. In PixelRL, each pixel has been assigned an agent so the number of agents is equivalent to the number of pixels, and it achieves excellent results compared to traditional methods. At the same time, another multi-agent algorithm is proposed in Ref.[9], where each agent controls a sliding square window. This method is more robust and has better generalization based on the results of the experiments on the CIFAR-10 dataset and the MNIST dataset [10].

2. LITERATURE SURVEY

Wang, et al. [11] proposed a fusion-enhanced attention mechanism and efficient self-architecture lightweight YOLO (SLGA-YOLO) to overcome the existing target detection algorithms' poor computational efficiency and low defect-detection accuracy. The integration of simplified attention module (SimAM) and Large Separable Kernel Attention (LSKA) fusion strengthens the attention mechanism, improving the detection performance, while significantly reducing computational complexity and memory usage. To enhance the generalization ability of the model's feature extraction, we replaced part of the basic convolutional blocks with the self-designed GhostConvML (GCML) module, based on the addition of p2 detection. Chen, et al. [12] developed an automated optical inspection (AOI) system to address these challenges. The system integrates a semantic segmentation neural network to handle external surface detection and an anomaly detection model to detect internal defects. In terms of internal defect detection, the GC-AD-Local model we tested achieved 100% accuracy on experimental images, and the results were relatively stable.

Andriosopoulou, et al. [13] examines the applicability of popular deep neural network types, which are widely employed for object detection tasks, in recognizing surface defects of parts that are produced through a die-casting process. The data used to train the networks belong to two different datasets

consisting of images that contain various types of surface defects and for two different types of parts. The first dataset is freely available and concerns pump impellers, while the second dataset has been created during the present study and concerns an automotive part. For the first dataset, Faster R-CNN and YOLOv5 detection networks were employed yielding satisfactory detection of the various surface defects, with mean average precision (mAP) equal to 0.77 and 0.65, respectively.

Sadikoglu, et al. [14] implemented a light Convolutional Neural Network (CNN) model to detect casting product defects without pooling operation. The model is trained and tested using casting product images, and the evaluation is performed using the sensitivity, specificity, balanced accuracy, and F1 Score. The implemented CNN model achieved 99.11%, 99.61%, 99.45%, and 99.36% sensitivity.

Lin, et al. [15] proposed to simulate the difference operation of biological vision; the suspicious defect area is directly detected according to the gradient threshold in this stage. In the inter-frame learning stage, the convolution neural network is established based on deep learning strategy to extract defect features from a suspicious defect area; a deep learning feature vector is obtained in this stage. The similarity degree of the suspicious defect area is computed by a feature vector; a casting defect is tracked by the similarity matching of the suspicious defect in continuous frames; then, the false defects (such as noise) are excluded after defect tracking. The experimental results show that the false rate and missed rate for detection of casting defects are less than 4%, and the accuracy of the defect detection is more than 96%, which proves the robustness of the proposed method.

Rupesh Gupta, et al. [16] included the research of casting defect detection is analyzed, including surface defect and internal defect detection technology, as well as image

preprocessing, algorithm framework optimization and design innovation direction. By reviewing a variety of deep learning models and algorithms, such as CNN, transfer learning, lightweight models, and improved U-Net models, the potential of these techniques in improving the efficiency and accuracy of defect detection is demonstrated. In addition, this paper discusses the three key modules of visual defect detection technology: image acquisition, image preprocessing and image analysis, and compares the characteristics of one-stage and two-stage object detection algorithms.

Zhang, et al. [17] detected a transformer with a multi-scale fusion attention mechanism, considering comprehensive features. Firstly, a novel joint data augmentation (JDA) method is constructed to alleviate the imbalanced dataset issue by effectively increasing the number of sample data. Then, an attention-based channel-adaptive weighting (ACAW) feature enhancement module is established to fully apply complementary information among different feature channels, and further refine feature representations. Consequently, a multi-scale feature fusion (MFF) module is proposed to integrate high-dimensional semantic information and low-level representation features, enhancing multi-scale defect detection precision.

Kaufmann, et al. [18] created the data used for the model were collected from a medium-sized steel and cast iron foundry in which components ranging from 1 to 100 kg in weight are produced from wear and heat resistant cast iron and steel materials. This includes all production-relevant data from the melting and casting process, as well as from the mold production, the sand preparation and component quality related data from the quality management department.

3. PROPOSED SYSTEM

The proposed system presents a GUI-based machine learning system designed to classify

casting product defects from image data. The system integrates dataset handling, preprocessing, model training, evaluation, and prediction within a single interactive interface. Images are converted into numerical feature vectors through resizing, normalization, and flattening, enabling compatibility with classical machine learning algorithms. Multiple models such as MNB, DTC, and RFC are trained and evaluated using standard performance metrics. The system also includes visualization techniques for performance comparison and supports real-time image prediction. The use of model saving and data caching ensures efficient execution and reduces repeated computations as shown in Fig. 3.

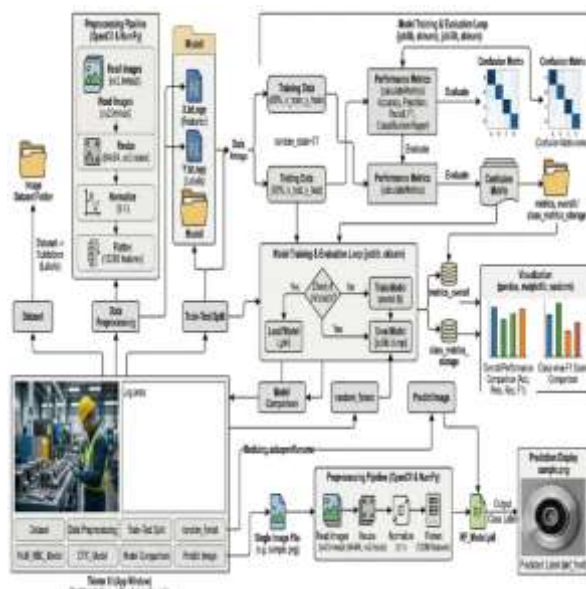


Fig. 3: System architecture for casting product defects from image data.

Dataset Upload and Category Extraction

The process begins with selecting a dataset folder using a file dialog interface. The system automatically identifies all subfolders, which act as class labels for classification. These categories are displayed in the interface to confirm successful dataset loading.

Image Preprocessing and Dataset Creation

Images are read from each category folder using OpenCV and resized to a fixed dimension of 64×64. Each image is normalized and flattened into a one-dimensional array. The processed data is stored in input and output arrays and saved as NumPy files for reuse.

Loading Pre-processed Data

The system checks whether pre-processed NumPy files already exist in the storage directory. If available, the arrays are loaded directly instead of reprocessing images. This step improves efficiency by reducing computation time during repeated executions.

Train-Test Data Splitting

The dataset is divided into training and testing subsets using an 80:20 ratio. A fixed random state ensures reproducibility of results. This separation allows the models to be trained on one portion and evaluated on unseen data.

Model Training and Saving

Multiple machine learning models, including MNB, DTC, and RFC, are trained using the training dataset. Each trained model is saved as a file for future use. If a saved model already exists, it is loaded directly instead of retraining.

Performance Evaluation and Metrics Calculation

The trained models are evaluated using accuracy, precision, recall, and F1-score. Class-wise metrics are also computed to analyze performance across different categories. Confusion matrices are generated and visualized for better understanding.

Performance Comparison Visualization

The system generates graphical comparisons of model performance using bar charts. Overall metrics and class-wise F1 scores are

visualized to identify the best-performing algorithm. Values are displayed on the graphs for clarity.

Image Prediction Using Trained Model

A new image is selected through the interface and pre-processed in the same manner as training data. The trained RFC model predicts the class label. The predicted result is displayed on the image for user interpretation.

GUI-Based Interaction Flow

All operations are controlled through a user-friendly graphical interface with buttons for each step. The interface provides real-time feedback and results in a text area. This design allows users to execute the entire workflow without writing code.

RFC model

RFC is an ensemble learning algorithm that combines multiple decision trees to improve performance as depicted in Fig. 4. In this study, it processes flattened image features to perform classification. Each tree is trained on random subsets of data and features. This randomness reduces overfitting and improves generalization. The final prediction is obtained using majority voting across all trees. It is robust to noise and handles high-dimensional data effectively. RFC provides higher accuracy compared to single models. It is used as the primary model for final prediction in this study.

Internal Operations

- Generates multiple bootstrap samples from the training dataset.
- Selects random subsets of features for each tree during training.
- Builds multiple decision trees independently on sampled data.
- Combines predictions from all trees using majority voting.

- Produces final output based on aggregated decision of all trees.

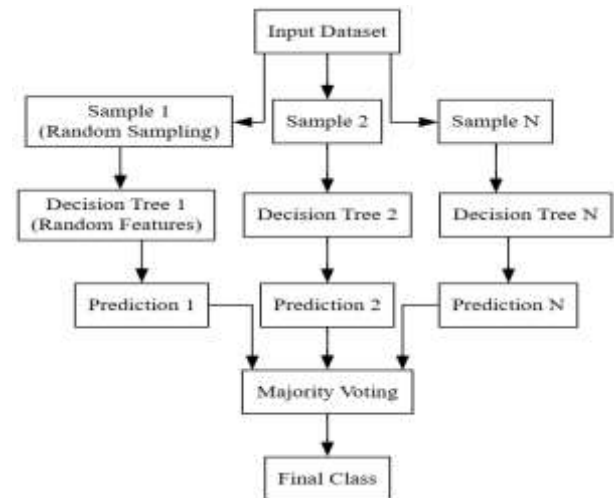


Fig 4: Internal Architecture of RFC Ensemble Model.

Advantages

- Reduces overfitting by combining multiple decision trees.
- Provides higher accuracy and better generalization.
- Handles large datasets and high-dimensional features effectively.
- Robust to noise and outliers in data.
- Supports feature importance estimation for better insights.
- Works efficiently for both classification and regression tasks.

4. Results Description

Fig. 5 shows the graphical user interface designed for executing the entire workflow. The interface contains multiple buttons such as Dataset, Data Preprocessing, Train-Test Split, and different model training options. It provides a structured environment for users to perform operations without coding. The layout is organized to guide the user step-by-step through the pipeline. The background design

reflects industrial inspection context. The text display area is used to show outputs and logs. This interface ensures ease of use and accessibility. Overall, it integrates all functionalities into a single platform.

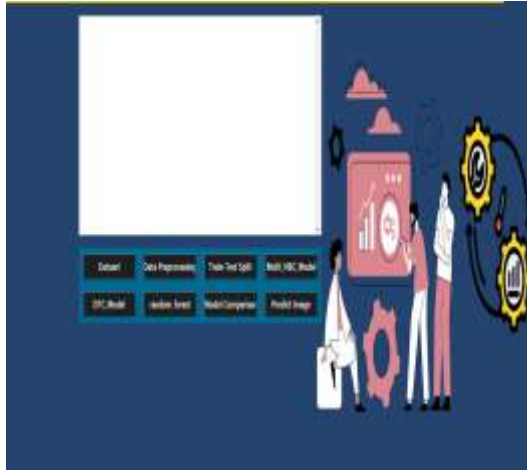


Fig. 5: GUI for defect detection in casting products

Fig. 6. shows the confusion matrix for the MNB model. It represents the classification performance across two categories: defect and non-defect. The matrix indicates correct and incorrect predictions made by the model. While the model performs reasonably well, there are noticeable misclassifications. The values show that some defect images are incorrectly predicted as non-defect. This affects overall precision and recall. However, the model still provides a baseline performance. It highlights the limitations of probabilistic methods on image data.

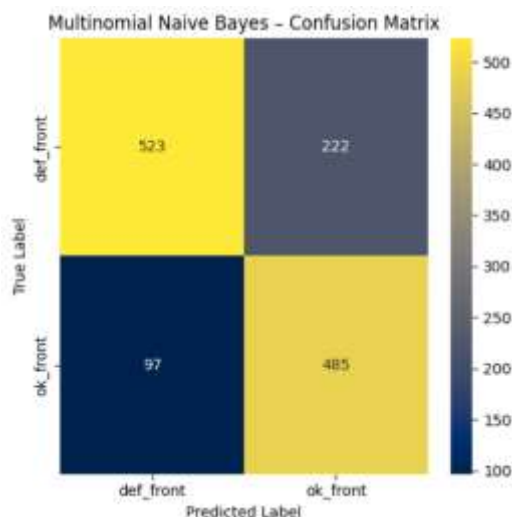


Fig. 6: Confusion matrix obtained using MNB Model

Fig. 7. illustrates the confusion matrix for the DTC. The model shows improved classification accuracy compared to Naive Bayes. Most defect and non-defect samples are correctly classified. The number of misclassifications is significantly reduced. The tree-based approach effectively captures feature patterns. This leads to better separation between categories. The confusion matrix reflects strong predictive capability. Overall, the DTC performs better than the baseline model.

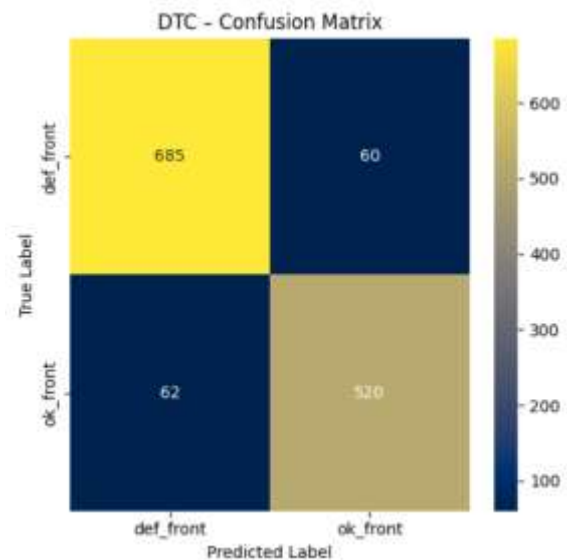


Fig. 7: Confusion matrix obtained using DTC

Fig. 8 represents the confusion matrix for the RFC model. It shows the highest accuracy among all models. The number of correct predictions is maximized while errors are minimized. The ensemble approach improves robustness and generalization. Very few samples are misclassified. This demonstrates the effectiveness of combining multiple decision trees. The model handles feature variability efficiently. RFC proves to be the best-performing model in this study.

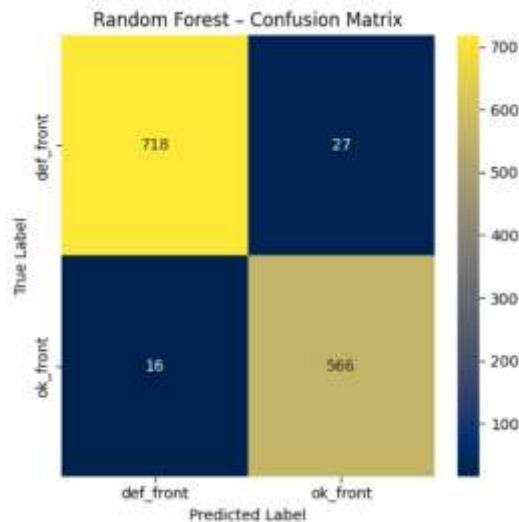


Fig. 8: Illustration of Confusion matrix obtained using RFC model.

Comparative Analysis

Fig. 9 compares the performance of all algorithms using metrics such as accuracy, precision, recall, and F1-score. The RFC model achieves the highest scores across all metrics. DTC also performs well but slightly below RFC. MNB shows comparatively lower performance. The graph provides a clear visual comparison of model effectiveness. It helps in selecting the best algorithm objectively. The results confirm the superiority of ensemble learning. This visualization supports final model selection.

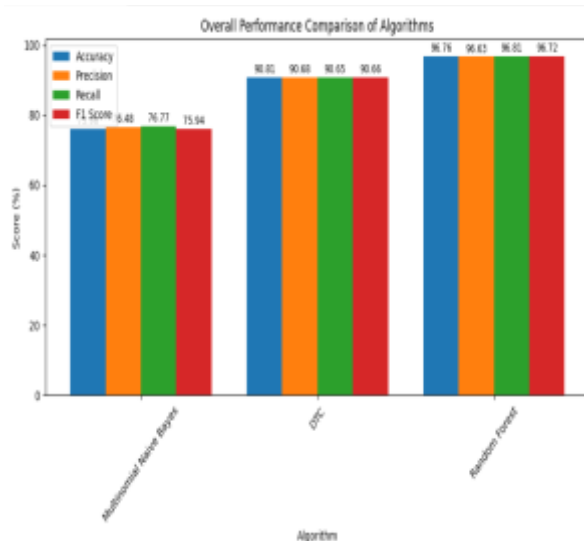


Fig. 9: Overall Performance Comparison Graph.

Fig. 10 presents the F1-score comparison for each class across different models. It shows how well each algorithm performs on individual categories. RFC achieves the highest F1-scores for both defect and non-defect classes. DTC also shows strong performance. Naive Bayes has lower scores due to misclassification. The graph highlights class-level performance differences. It ensures balanced evaluation across categories. This helps in understanding model reliability for each class.

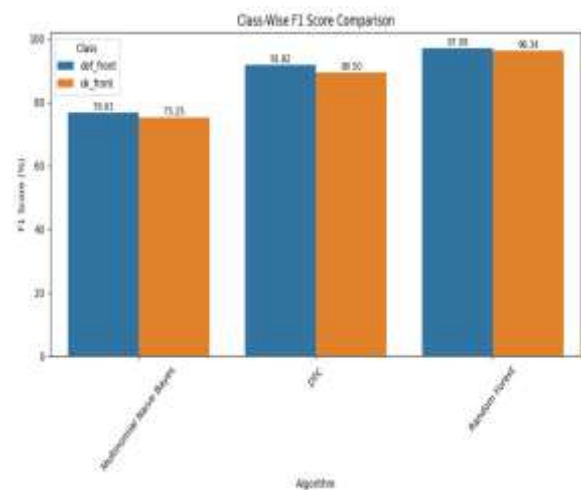


Fig. 10: Class-wise F1 Score Comparison.

Fig. 11. shows the prediction result for a test image. The system correctly identifies the image as "def_front". The output is displayed visually along with the input image. This confirms the successful working of the prediction pipeline. The model effectively processes unseen data. The result demonstrates real-time applicability of the system. The classification is accurate and clearly presented. This validates the effectiveness of the trained RFC model.

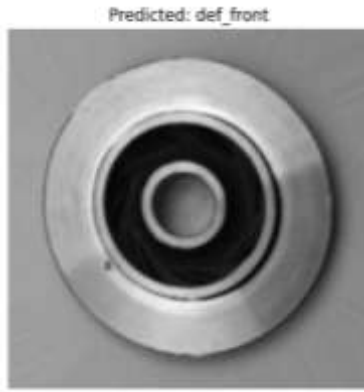


Fig. 11: Final Prediction Output.

This section presents a comparative analysis of the performance of different machine learning models used for defect classification in casting images. The evaluation focuses on three algorithms: MNB, DTC and RFC Classifier (RF). Each model is assessed using standard performance metrics such as accuracy, precision, recall, and F1-score. The comparison aims to identify the most effective algorithm for accurate defect detection. By analyzing both overall and class-wise performance, the study highlights the strengths and limitations of each model. This analysis supports the selection of the best-performing model for real-world industrial applications.

Performance Comparison

Table 1 clearly shows significant differences among the three models. MNB achieves an accuracy of 75.96%, indicating moderate performance in classification tasks. Its probabilistic nature makes it less effective for image-based data, resulting in higher misclassification rates. The class-wise metrics reveal imbalance, where one class achieves higher recall while the other suffers from lower precision. DTC improves performance significantly, achieving an accuracy of 90.81%. It effectively captures feature relationships and reduces classification errors. The class-wise metrics for DTC are well balanced, showing consistent performance across both categories. However, DTCs may still suffer from overfitting in certain cases.

Table 1: Performance Comparison of Machine Learning Models

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
MNB	75.96	76.48	76.77	75.94
DTC	90.81	90.68	90.65	90.66
RFC	96.76	96.63	96.81	96.72

RFC outperforms all other models with an accuracy of 96.76%. The ensemble approach combines multiple decision trees, leading to improved generalization and reduced variance. The precision, recall, and F1-score values are consistently high, indicating strong classification capability. Class-wise performance also shows balanced and superior results for both defect and non-defect categories. RFC achieves the highest F1-scores, demonstrating its ability to handle both classes effectively. The improvement over DTC highlights the advantage of ensemble learning. The comparison confirms that RFC is the most reliable and accurate model for this classification task.

5. CONCLUSION

This study successfully developed and evaluated an automated image-based casting defect classification system using machine learning techniques. A comprehensive framework consisting of image preprocessing, feature extraction, data balancing, model training, and performance evaluation was implemented to ensure accurate and reliable defect detection. The comparative analysis of Multinomial Naive Bayes (MNB), Decision Tree Classifier (DTC), and Random Forest Classifier (RFC) demonstrated that ensemble learning techniques significantly outperform traditional single-model approaches. Among all evaluated models, the Random Forest Classifier achieved the best performance in terms of accuracy, precision, recall, and F1-score, highlighting its effectiveness in identifying complex defect patterns within

casting images. The application of image normalization, feature vector generation, and SMOTE-based class balancing contributed substantially to improving model robustness and reducing classification bias. Furthermore, confusion matrix analysis and performance visualizations confirmed the consistency and reliability of the proposed approach. The developed system minimizes the limitations of manual inspection by providing faster, more accurate, and consistent defect detection. Its ability to automate the quality inspection process makes it highly suitable for modern industrial environments where efficiency, scalability, and reliability are essential. Overall, the proposed machine learning-based framework offers a practical and cost-effective solution for casting surface defect classification, contributing to enhanced quality control, reduced inspection time, and improved manufacturing productivity.

REFERENCES

- [1] Jiang, X.; Feng Wang, X.; Chen, D. Research on Defect Detection of Castings Based on Deep Residual Network. In Proceedings of the 2018 IEEE 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 13–15 October 2018; pp. 1–6.
- [2] Du, W.; Shen, H.; Fu, J.; Zhang, G.; He, Q. Approaches for improvement of the X-ray image defect detection of automobile casting aluminum parts based on deep learning. *NDT Int.* **2019**, *107*, 102144.
- [3] Duan, L.; Yang, K.; Ruan, L. Research on automatic recognition of casting defects based on deep learning. *IEEE Access* **2020**, *9*, 12209–12216.
- [4] Ferguson, M.; Ak, R.; Lee, Y.T.T.; Law, K.H. Automatic localization of casting defects with convolutional neural networks. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1726–1735.
- [5] Lee, J.H.; Oh, H.M.; Kim, M.Y. Deep learning based 3D defect detection system using photometric stereo illumination. In Proceedings of the 2019 IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Okinawa, Japan, 11–13 February 2019; pp. 484–487.
- [6] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2012; Volume 25.
- [7] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
- [8] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- [9] Chen, L.; Li, S.; Bai, Q.; Yang, J.; Jiang, S.; Miao, Y. Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sens.* **2021**, *13*, 4712.
- [10] Rawat, W.; Wang, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.* **2017**, *29*, 2352–2449.

