

# A Lightweight Defect Classification Method for Latex Gloves Based on Image Enhancement

Yong Ren<sup>1,\*</sup>, Dong Liu<sup>1</sup> and Sanhong Gu<sup>2</sup>

<sup>1</sup> Applied Technology College Of Soochow University  
215325 SuZhou, China  
renyong@suda.edu.cn  
flovvv@163.com

<sup>2</sup> Suzhou Dechuang Measurement and Control Technology Co., Ltd.  
215104 Suzhou, China  
threehym@163.com

**Abstract.** This paper presents a glove defect classification method that integrates image enhancement techniques with a lightweight model to enhance the efficiency and accuracy of glove defect classification in industrial manufacturing. A dataset comprising images of five types of gloves was collected, totaling 360 sample images, for the training and validation of a deep learning-based glove defect classification model. Image enhancement techniques, including super-pixels, exposure adjustment, blurring, and limited contrast adaptive histogram equalization, increased dataset diversity and size, improving model generalization. Based on the lightweight model MobileNetV2, the model was improved by reducing the number of input image channels through grayscale conversion and optimizing the loss function. Experimental results demonstrate that the improved MobileNetV2 model achieved an average accuracy of 97.85% on both the original and enhanced datasets, effectively mitigated overfitting phenomena, and exhibited a significantly faster training speed compared to the ResNet34 and ResNet50 models.

**Keywords:** glove defect classification, machine vision, image enhancement, deep learning, lightweight model, mobilenetv2.

## 1. Introduction

Industrial defect classification is an important technology to ensure product quality. The rapid development of new technologies in the fields of machine vision, image processing, and deep learning has also driven the significant progress of industrial defect classification technology. Currently, it can be applied to the detection of various industrial products such as metals, semiconductors, textiles, and rubber [21].

The limitations of traditional manual inspection methods, such as low efficiency, high cost, and strong subjectivity, can no longer meet the needs of modern industrial production. In the field of latex glove production, with the advancement of industrial automation and intelligence, glove defect classification technology is also gradually transitioning from traditional manual visual inspection to automated detection based on machine vision, which makes the machine vision-based defect classification technology play an

---

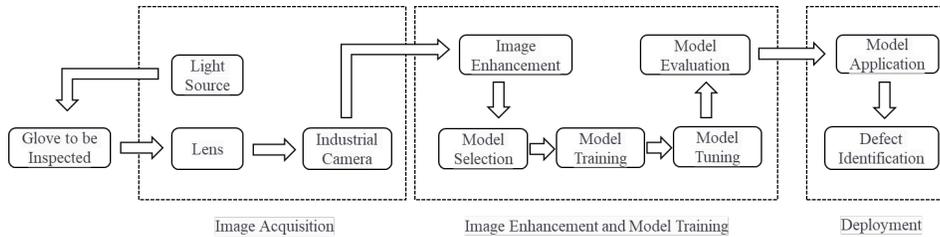
\* Corresponding Author

increasingly important role in product quality assurance. Therefore, the research and development of efficient defect classification technology has become particularly urgent [20].

Traditional machine vision-based defect classification methods mainly relied on image processing technologies, such as edge detection and template matching [3]. These methods performed well in handling simple or rule-based industrial images, but their performance was limited in complex scenarios. With the application of machine learning technology, researchers have begun to explore statistical learning-based methods, such as support vector machines (SVMs) and random forests, which can better handle the local features and classification problems of images [10, 24].

The rise of deep learning technology has brought new development opportunities for defect classification. Convolutional neural networks (CNNs) have been widely applied to defect classification tasks, especially in image classification, object detection, and semantic segmentation [6, 16, 30].

The main goal of glove defect classification in the industrial production process is to quickly distinguish the types of qualified and nonqualified products, so as to control product quality. Due to the variety of gloves and defect types, collecting and labeling a large amount of high-quality data is costly and challenging [27], and there are still problems such as strong data dependency and weak model generalization ability. In addition, real-time and computational resource requirements are also limiting factors. Therefore, this paper proposes a lightweight latex glove defect classification method based on image enhancement.



**Fig. 1.** The process of glove defect classification based on machine vision

## 2. Related Work

### 2.1. Defect classification based on machine vision

The typical process of defect classification based on machine vision usually includes three parts: image acquisition, image enhancement and model training, and deployment application.

Firstly, an industrial camera is used to capture high-quality images of products. The collected data is then subjected to image enhancement. Next, a suitable model is selected for training. Finally, the trained model is applied to perform defect classification. For example, The overall process of glove defect classification is illustrated in Figure 1.

Early defect detection methods relied heavily on manual inspection and basic image processing techniques. While early methods can be effective in specific scenarios, they are labor-intensive and do not scale well with increasing production volumes. To overcome these limitations, conventional image processing techniques such as thresholding, edge detection, and template matching were introduced [15]. Thresholding involves setting a gray-level threshold to segment images into foreground and background, facilitating defect identification based on pixel intensity differences [11]. Edge detection algorithms, such as Canny and Sobel, are designed to identify abrupt changes in image brightness, thereby outlining object boundaries that can indicate the presence of cracks or scratches [19]. Template matching compares a region of interest in the test image with predefined templates, allowing for the detection of deviations that signify defects [32]. Despite their simplicity and high computational efficiency, traditional methods often face challenges when dealing with complex defect patterns, varying lighting conditions, and noisy images. This has prompted the exploration of more advanced techniques, particularly those that leverage deep learning.

In recent years, deep learning has transformed defect detection by automating feature extraction and enabling the learning of intricate patterns from large datasets. Convolutional Neural Networks (CNNs) have become foundational in deep learning-based defect detection due to their capability to hierarchically learn spatial hierarchies of features from image data [6]. CNNs utilize convolutional layers that apply learnable filters to the input images, allowing them to automatically capture essential features such as edges, textures, and shapes at multiple scales. This hierarchical feature extraction process means that initial layers may learn simple patterns like edges and colors, while deeper layers can recognize more complex structures, such as patterns specific to defects [7]. Furthermore, the architecture of CNNs can be tailored to specific defect detection tasks by adjusting parameters such as the number of layers, filter sizes, and activation functions. This flexibility allows researchers to optimize CNNs for particular types of defects or image characteristics, thereby improving classification accuracy [12].

Beyond CNNs, other deep learning architectures have also been employed for defect detection. Generative Adversarial Networks (GANs), for instance, are effective for anomaly detection by training a generator to create realistic, defect-free images, while a discriminator distinguishes between real and generated images. This setup allows GANs to flag anomalies as deviations from the learned distribution [29]. Similarly, Autoencoders can be adapted for defect detection by reconstructing input images with minimal reconstruction error. Significant discrepancies between the original and reconstructed images can indicate potential defects, making Autoencoders valuable for identifying anomalies that may not be well-represented in the training data [14].

## 2.2. Image enhancement techniques

Image enhancement techniques play a pivotal role in enhancing the efficiency and generalization capabilities of deep learning models, especially when dealing with limited sample data [25, 8]. These techniques serve to expand the diversity and size of the dataset by simulating various changes that may occur in the real world and generating new image samples by applying a series of transformations to the original image [18].

One common approach to image enhancement is data augmentation, which involves applying a series of transformations to the original images, such as rotation, scaling, flip-

ping, and color adjustments. These transformations help the model generalize better by exposing it to a wider range of visual patterns during training. In addition to data augmentation, other image enhancement techniques, such as histogram equalization and noise reduction, can also be employed to further improve image quality and feature extraction [1, 13]. These techniques are particularly useful in scenarios where the original images are of low quality or contain significant noise. Recent advancements in generative adversarial networks (GANs) have also opened up new possibilities for image enhancement [28]. GANs can generate highly realistic synthetic images that can be used to augment the training dataset, further improving the performance of deep learning models in glove defect classification tasks.

In summary, the integration of image enhancement techniques with deep learning models has significantly advanced the field of glove defect classification. By leveraging these techniques, researchers have been able to improve classification accuracy, reduce manual inspection efforts, and enhance the robustness and generalization capabilities of their models.

### 3. Materials and Methods

#### 3.1. Dataset construction

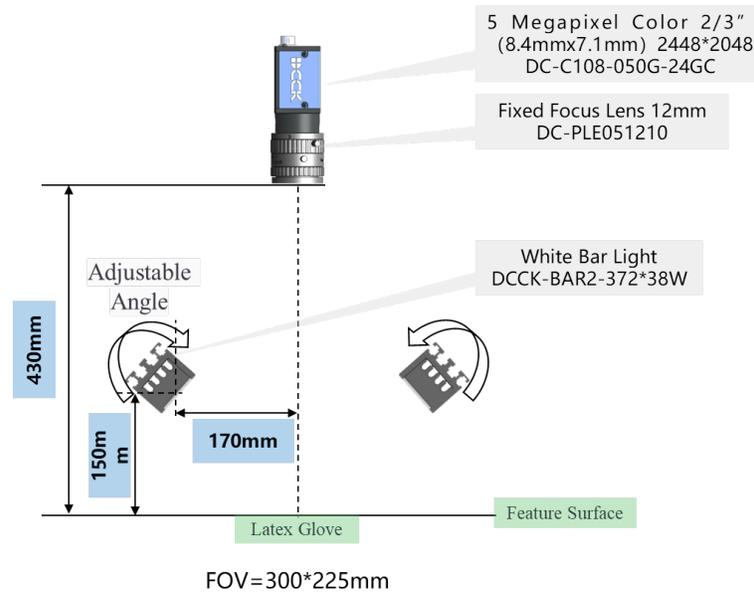
The choice and setup of the lens and industrial camera are crucial for the quality of captured images [5]. In this study, a 5 megapixel color camera and a 12mm fixed focus lens of Detron were selected, and white stripe light and black background were used for shooting. The side view of the camera is shown in Figure 2.

In this paper, a dataset containing five types of glove images, totaling 360 sample images, was collected to train and validate a glove defect classification model based on deep learning. The dataset is crucial for effectively training the model to recognize various types of defects that may occur in latex gloves, which are commonly used in industrial settings. Each type of image reflects different defect characteristics, as shown in Table 1.

The inclusion of these diverse image types allows for a comprehensive evaluation of the model's ability to detect defects that can occur in real-world scenarios. For instance, minor damage such as small cracks or wear may not render the gloves immediately unusable but could lead to issues over time, affecting the safety of users. On the other hand, major damage, including significant tears, directly impacts the gloves' usability and poses a safety risk.

To address the challenges presented by varying defect types and conditions, the following image enhancement techniques are introduced in this study: color dithering, which can simulate the effects of different lighting conditions and increase the color diversity of the dataset by randomly changing the color values in the image; brightness adjustment, which can better adapt to different ambient lighting situations by adjusting the brightness of the image so that the model can deal with a variety of scenarios ranging from bright to dim; and contrast adjustment, which enhances or reduces the contrast of the image that can help the model learn to recognize defects under different contrast conditions.

Additionally, other methods such as superpixels, blur, solarize, affine transformations, Canny edge detection, and CLAHE are utilized for image enhancement, as illustrated in Figure 3. The specific image augmentation methods and their parameters are summarized in Table 2.



**Fig. 2.** Camera Side View and Parameter Description

In order to compare and analyze the impact of image enhancement on the model training effect, the dataset (Raw) is the original glove sample image set, without any image enhancement processing. The dataset (Aug) adopts 6 transformation methods for image enhancement. The above image enhancement combination builds the training dataset as shown in Table 3.

### 3.2. Network model structure

In this study, lightweight network MobileNetV2 was selected as the basic model architecture, and the classic ResNet34 and ResNet50 network models were used in the comparative experiment. ResNet networks [4] effectively solve the problem of gradient disappearance in deep neural networks by introducing residual connections, thus allowing the network to train deeper model structures. The difference between ResNet34 and ResNet50 lies not only in their depth and number of layers, but also in their construction of Residual blocks. ResNet34 uses the Residual Block, while ResNet50 uses a Bottleneck block, as shown in Figure 4.

ResNet34 includes a 34-layer network structure with fewer parameters, making it more efficient when computing resources are limited or on small datasets. ResNet50, which contains a 50-layer network structure, has more parameters and more complex feature extraction capabilities, and is suitable for large datasets and complex image recognition tasks. The MobileNetV2 [22] network employs an Inverted Residual Block structure (as shown in Figure 5) and Linear Bottlenecks, enhancing model efficiency and accuracy while maintaining low computational cost and minimal memory usage.

**Table 1.** Glove Image Types and Sample Quantities

Type	Quantity(Pieces)	Description
Normal Images (Ok)	200	Glove images without visible defects
Empty Images (Empty)	18	Blank images with no gloves captured
Minor Damage Images (Minor)	27	Glove images with small cracks or wear
Major Damage Images (Major)	72	Glove images with significant damage or tears
Dirty Images (Dirty)	43	Glove images contaminated with oil stains or similar substances

To reduce the computational complexity of the model, this study converts input sample images to grayscale, removing color information and retaining only luminance data, which means the model is trained using single-channel images. Consequently, the first input layer parameter of the MobileNetV2 model is changed from  $224 \times 224 \times 3$  to  $224 \times 224 \times 1$ . The overall structure of the adjusted MobileNetV2 network is shown in Table 4, where  $t$  represents the expansion factor,  $c$  denotes the depth of the output feature map (channel),  $n$  indicates the number of repetitions of the bottleneck, and  $s$  refers to the stride.

Meanwhile, ReLU6 is used as the activation function, a Rectified Linear Unit activation function [23] with a rectified linear unit output limit of 0 to 6, which can be mathematically defined as:

$$ReLU6(x) = \min(\max(0, x), 6) \quad (1)$$

### 3.3. Loss Function Optimization

In addition, to train the model to accurately identify the defect type of gloves, the Loss function employs Multi-Class Cross-Entropy loss, which measures the difference between the probability distribution predicted by the model and the true label [31]. For each sample, the loss function can be expressed as:

$$L = \sum_{i=1}^C y_i \log(P_i) \quad (2)$$

Where:  $C$  is the number of categories,  $y_i$  is the unique thermal coding of the real label, and  $p_i$  is the probability distribution predicted by the model.

To improve model performance, Mini-batch Gradient Descent is utilized for training, specifically using the Adam optimizer [17, 26]. The Adam optimizer combines the benefits of Momentum and adaptive learning rates, allowing for automatic adjustment of the learning rate for each parameter, thus speeding up the training process and increasing convergence speed.

**Table 2.** Image Augmentation Methods

Method	Parameter Settings	Operational Steps
Color Dithering	RGB values within $\pm 10\%$	Use <code>PIL</code> to modify pixel values randomly.
Brightness Adjustment	Range: [0.5, 1.5]	Use <code>ImageEnhance.Brightness</code> to adjust brightness.
Contrast Adjustment	Range: [0.5, 1.5]	Use <code>ImageEnhance.Contrast</code> to adjust contrast.
Superpixels	N/A	Apply superpixel segmentation using <code>skimage</code> .
Blur	Adjustable kernel size	Use Gaussian blur to reduce noise and details.
Solarize	Threshold parameter	Invert brightness of pixels above a certain threshold.
Affine Transform	Rotation angle, scaling factors	Apply transformations to create diverse perspectives.
Canny Edge Detection	N/A	Highlight edges using the Canny algorithm.
CLAHE	Clip limit, grid size	Enhance local contrast using CLAHE algorithm.

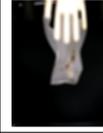
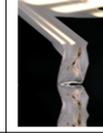
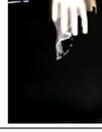
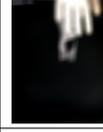
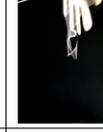
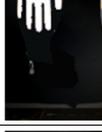
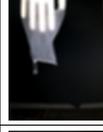
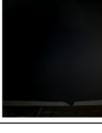
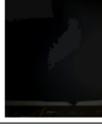
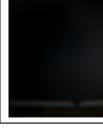
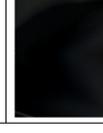
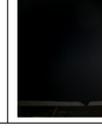
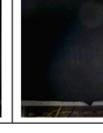
**Table 3.** Image Enhancement Combination for Constructing the Training Dataset

Dataset	Sample Classification Quantity	Total(Pieces)
Original Dtaset(Raw)	Normal (OK) 200, Empty 18, Minor 27, Major 72, Dirty 43	360
Augmented Dtaset(Aug)	Normal (OK) 200, Empty 108, Minor 162, Major 432, Dirty 253	1155

On the basis of multi-class cross-entropy Loss function, weighted loss function and Focal Loss are introduced. The weighted loss function increases the model's attention to a few categories by assigning different weights to each category. The weights for each class can be set based on the frequency of samples in the dataset. For example:

$$\text{Weight}_i = \frac{N}{n_i}$$

Where:  $N$  is the total number of samples, and  $n_i$  is the number of samples in class  $i$ . By assigning higher weights to minority classes (e.g., Minor damage), the model focuses more on these samples during training. The weights are fine-tuned through cross-validation to find a balance that maximizes overall model accuracy.

	Initial	Augmented					
		Superpixels	Blur	Solarize	Affine	Canny	CLAHE
OK							
Dirty							
Major							
Minor							
Empty							

**Fig. 3.** Example of Image Enhancement

The focus loss makes the model more focused on the learning of difficult classification samples by reducing the loss weights of easy classification samples. Specifically, focus loss can be expressed as

$$FL(p_i) = -\alpha_t(1 - p_t)^\gamma(p_t) \quad (3)$$

In this formula,  $\alpha_t$  is the class weight,  $p_t$  is the prediction probability of the model for the correct class, and  $\gamma$  is the regulator that adjusts the rate at which easy samples are down-weighted. A typical value for  $\gamma$  might be set to 2, but it can be optimized through grid search or random search techniques to determine the value that yields the best validation accuracy.

By introducing an improved loss function, the model can deal with class imbalance more effectively, thus improving the overall detection accuracy.



**Fig. 4.** Two Types of Residual Blocks in ResNet

**Table 4.** The overall structure of the adjusted MobileNetV2

Input	Operator	t	c	n	s
$224^2 \times 1$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d $1 \times 1$	-	1280	1	1
$7^2 \times 1280$	avgpool $7 \times 7$	-	-	1	-
$1 \times 1 \times 1280$	conv2d $1 \times 1$	-	-	-	-

## 4. Experiment and Results

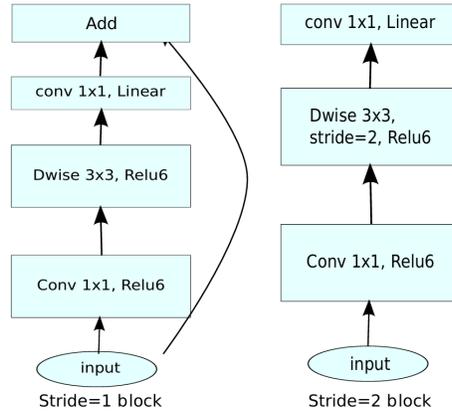
### 4.1. Experimental environment and model training

The experiments in this paper were conducted on a cloud server equipped with an NVIDIA GPU to ensure sufficient computing power for deep learning model training. The hardware specifications are as follows:

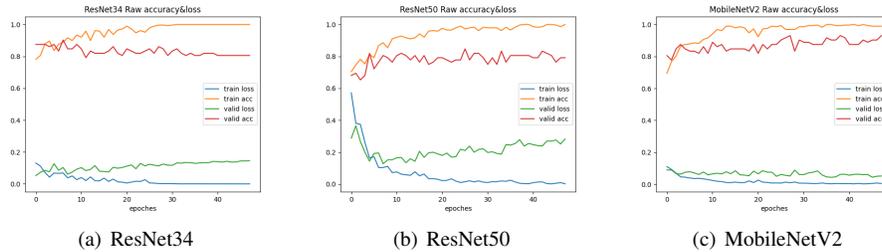
GPU: NVIDIA RTX A4000 with 16GB of video memory  
 CPU: 8 × Intel(R) Xeon(R) CPU E5-2686 v4 @ 2.30GHz  
 Memory: 60GB  
 Storage: 200GB

The software environment used for the experiments includes Ubuntu 20.04, Python 3.10, PyTorch 2.0.1, and CUDA 11.8. All training and testing experiments were performed in this environment.

Three models were trained: ResNet34, ResNet50 and MobileNetV2. Each model was trained for 50 epochs on both the Raw dataset and the Aug dataset. The initial learning rate for all gradient descent algorithms was set to  $1e-3$ , and the model training batch size was set to 4.



**Fig. 5.** Convolution block of MobileNetV2

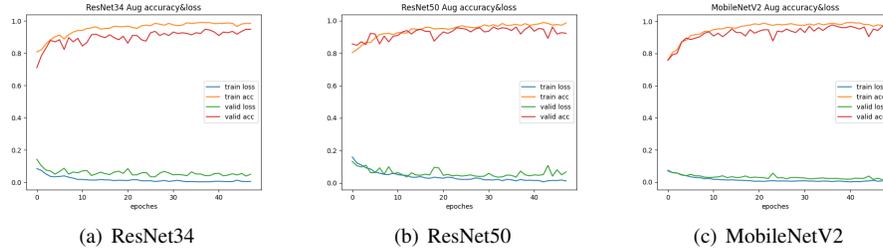


**Fig. 6.** Model Training Process on the Raw Dataset

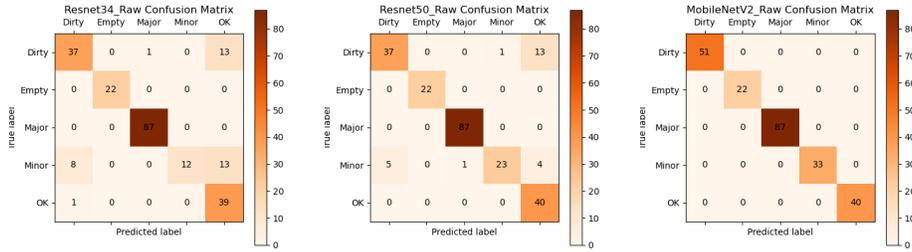
As shown in Figure 6, it can be observed that on the original dataset (Raw), both the ResNet34 and ResNet50 models begin to overfit after 10 training epochs due to insufficient training samples. The ResNet50 model, having a stronger feature fitting capability, tends to overfit more easily, resulting in a higher validation loss. In contrast, Figure 7 indicates that on the augmented dataset (Aug), overfitting for both ResNet34 and ResNet50 occurs only after 40 epochs, with a noticeable improvement in both training and validation accuracy, and a loss value approaching zero. The MobileNetV2 model outperformed the other two models both before and after enhancement.

**4.2. Confusion matrix, ROC curve and multi-classification evaluation index**

In order to ensure the effectiveness and generalization performance of the experimental results, the enhanced dataset is divided, and 20% (233 images) are set aside as the test dataset, and the original proportion of various types of images is maintained as far as possible. For the model trained by ResNet34, ResNet50 and MobileNetV2 on the Raw dataset, the confusion matrix and ROC curve of the test results [9] are shown in Figure 8 and Figure 9.



**Fig. 7.** Model Training Process on the Augmented Dataset



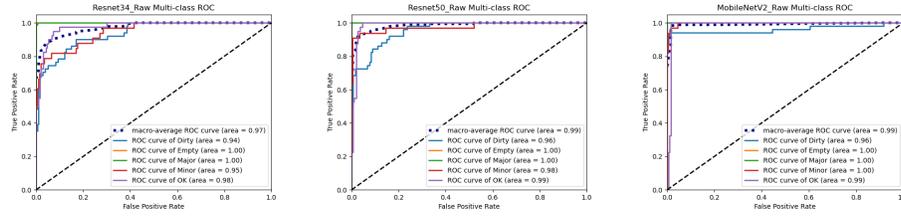
**Fig. 8.** Confusion Matrix of the Model on the Raw Dataset

According to the above test results, the prediction accuracy of the Minor image model is the lowest for all kinds of models, because the number of samples is relatively small and the difference between the minor image and normal image (OK) and Dirty image is not obvious. In the same case, although the Empty image has the smallest sample number, it has the most significant difference from other types of images, so it is easier to distinguish and has the highest prediction accuracy. However, due to the large number of samples and great difference from other types of images, Major image has the second highest prediction accuracy.

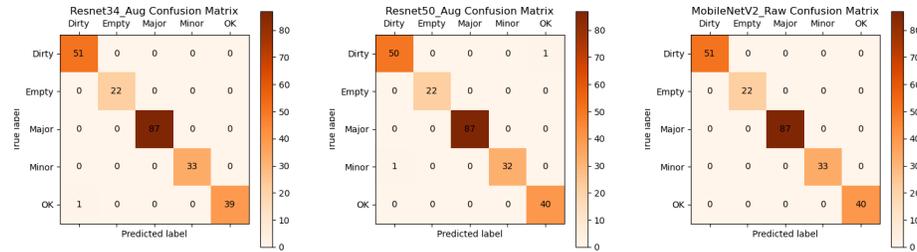
Training results of ResNet34, ResNet50, and MobileNetV2 models on the enhanced dataset (Aug). The confusion matrix and ROC curve tested are shown in Figure 10 and Figure 11.

It can be seen that in the enhanced dataset, the model training effect has been significantly improved, and all evaluation indicators of ResNet34 and ResNet50 models have exceeded 99%, and MobileNetV2 even reached 100%. Moreover, the Area Under Curve (AUC) of the five types of ROC curves almost reached 1, indicating good model performance.

This improvement can be attributed to increased dataset diversity from various augmentation techniques, which enable the model to learn robust features. Additionally, the use of enhanced data mitigates overfitting by encouraging the model to focus on essential features rather than memorizing specific examples. Finally, the introduction of challeng-



**Fig. 9.** ROC Curve of the Model on the Raw Dataset



**Fig. 10.** Confusion Matrix of the Model on the Augmented Dataset

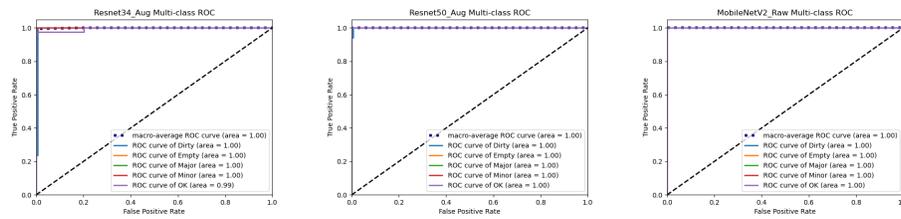
ing examples allows the model to better recognize subtle differences in glove defects, which is crucial for precision in industrial applications. These factors collectively demonstrate the effectiveness of image enhancement techniques in boosting detection accuracy and model reliability in glove defect classification tasks.

In addition, based on the confusion matrix of model test, Accuracy, Precision, Recall and F1 Score were calculated in this paper as evaluation indicators of model performance [2], and the experimental results were shown in Table 5.

**Table 5.** Model Evaluation Metrics Statistics

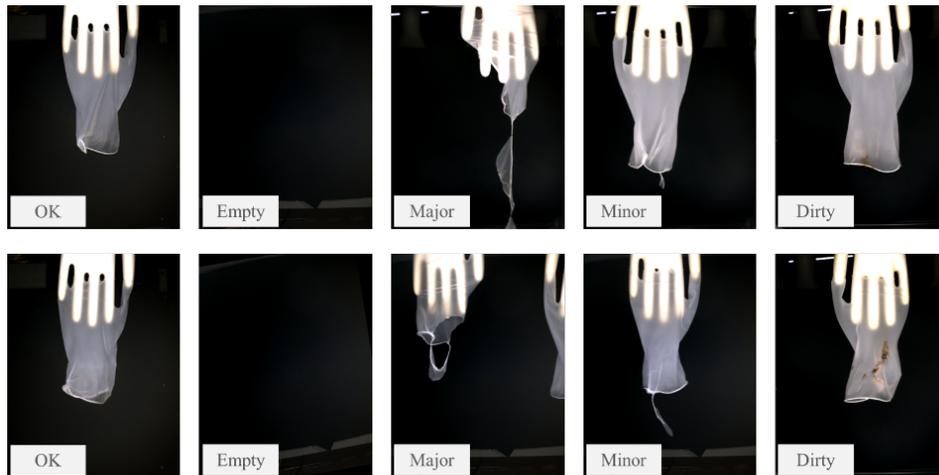
Evaluation Metrics	Resnet34		Resnet50		MobileNetV2	
	Raw	Aug	Raw	Aug	Raw	Aug
Accuracy	84.55	99.57	89.70	99.14	95.71	100
Precision	87.86	99.62	90.59	99.12	95.20	100
Recall	81.28	99.50	88.45	99.00	94.79	100
F1 Score	80.67	99.55	88.43	99.05	94.86	100

Comparative analysis of the experimental results shows that: In the training results on the original dataset, the four evaluation indexes of Resnet34 and Resnet50 models are



**Fig. 11.** ROC Curve of the Model on the Raw Dataset

between 80% and 90%, while the MobileNetV2 model is about 95%. The training results on the enhanced dataset showed that the four evaluation indexes of the three models were improved, with the Resnet34 and Resnet50 models exceeding 99%, and the improved MobileNetV2 model even reaching 100%. This shows that the combined image enhancement transform proposed in this paper can play a good role in improving the training effect of the model, and the training effect of the improved lightweight model MobileNetV2 is obviously better than that of the other two models. The glove classification detection effect under the optimal weight condition is shown in Figure 12.



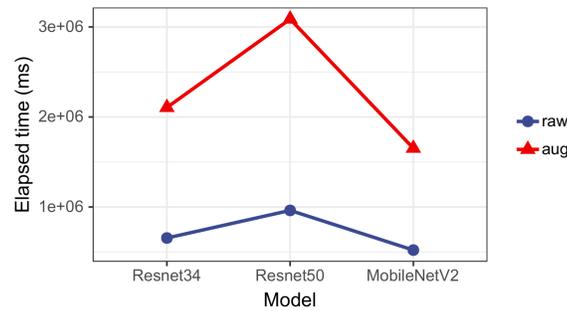
**Fig. 12.** Improved glove detection effect based on MobileNetV2 model

To understand the performance differences among the models, we examined key factors such as architecture, parameter count, computational complexity, and applicable scenarios. MobileNetV2 employs depthwise separable convolutions, reducing parameters and computational complexity compared to ResNet34 and ResNet50, enhancing generalization especially with limited data. With only 3.4 million parameters, MobileNetV2 offers faster training and less risk of overfitting, making it suitable for resource-constrained

environments and real-time applications. In contrast, ResNet34 and ResNet50, with approximately 21 and 25 million parameters respectively, require more computational resources but offer higher accuracy, suitable for scenarios where precision is prioritized and resources are abundant.

### 4.3. Model training time overhead

Under the same experimental conditions, Resnet34, Resnet50 and MobileNetV2 models do not use pre-training weights, and their respective training time costs on different datasets as shown in Figure 13.



**Fig. 13.** Comparison of Model Training Time Costs

The experimental results indicate that the training time of the model is primarily determined by factors such as the complexity of the model structure, the number of parameters, and the size of the training dataset. Compared to the ResNet34 and ResNet50 models, the MobileNetV2 model employs depthwise separable convolutions, effectively reducing both the parameter count and computational load. This allows it to maintain high performance while offering better computational efficiency, making it suitable for resource-constrained situations.

In conclusion, the introduction of image augmentation techniques has significantly enhanced the model's performance. By simulating different perspectives and scales, image augmentation increases the diversity of the dataset, enabling the model to learn more robust feature representations. Considering the performance requirements for glove image defect classification and the trade-off between model effectiveness and time expenditure, the MobileNetV2 model trained on the augmented dataset(Aug), demonstrates superior overall performance compared to the ResNet34 and ResNet50 models.

## 5. Conclusion and Discussion

Aiming at glove defect classification in industrial production, a deep learning method combining image enhancement technology and lightweight model is proposed in this paper. Three network models, ResNet34, ResNet50 and MobileNetV2, were used to compare and analyze the model training effect of the original dataset and the enhanced dataset.

The experimental results demonstrate that high-quality image samples, captured using machine vision, effectively simulate the diversity of real-world scenarios when subjected to image enhancement techniques. This approach significantly improves the model's generalization ability and detection accuracy, allowing it to identify a wider range of defects with greater reliability. Specifically, the enhancements not only facilitate better recognition of subtle defects but also enable the model to perform well under varying conditions that might not be present in the original dataset. Furthermore, the lightweight MobileNetV2 model, in particular, showcases a remarkable reduction in the number of parameters and computational complexity, making it suitable for real-time applications in industrial settings. By meeting the stringent performance requirements of practical industrial applications, this method provides a novel and efficient solution for glove defect classification. Ultimately, it contributes to improving the efficiency and overall product quality in industrial production processes.

Future researches on glove defect classification can focus on exploring more efficient image enhancement techniques and deep learning models, tailored to the needs of practical application scenarios. While our study has primarily employed some augmentation methods, such as brightness and contrast adjustments, it is still potential to investigate more advanced techniques, including Generative Adversarial Networks (GANs), which could generate diverse and realistic images to augment the dataset and enhance model generalization.

Furthermore, expanding the scope of research to more complex scenarios, including varying lighting conditions, occlusions, and diverse glove materials, is essential for assessing the model's adaptability in real-world applications. By addressing these challenges, we aim to continuously optimize and improve the accuracy and efficiency of glove defect classification systems, ensuring their robustness and applicability in diverse industrial settings.

Future studies could also consider the integration of multimodal data, such as thermal imaging or depth information, to enhance detection accuracy and reliability.

**Acknowledgments.** Teaching research project of Basic Computer Education in national colleges and Universities (2022-AFCEC-121); Ministry of Education Industry-University Collaborative Education Project (230804471307131).

## References

1. Acharya, U.K., Kumar, S.: Genetic algorithm based adaptive histogram equalization (gaahe) technique for medical image enhancement. *Optik* 230, 166273 (2021)
2. Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V.: Randaugment: Practical automated data augmentation with a reduced search space. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 702–703 (2020)
3. Gaidhane, V.H., Rani, A., Singh, V., et al.: An improved edge detection approach and its application in defect detection. In: *IOP conference series: materials science and engineering*, vol. 244, p. 012017. IOP Publishing (2017)
4. 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*, pp. 770–778 (2016)
5. How, Y.C., Nasir, A.F.A., Muhammad, K.F., Majeed, A.P.A., Razman, M.A.M., Zakaria, M.A.: Glove defect detection via yolo v5. *Mekatronika: Journal of Intelligent Manufacturing and Mechatronics* 3(2), 25–30 (2021)

6. Jha, S.B., Babiceanu, R.F.: Deep cnn-based visual defect detection: Survey of current literature. *Computers in Industry* 148, 103911 (2023)
7. Khanam, R., Hussain, M., Hill, R., Allen, P.: A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access* (2024)
8. Kim, J.S., Kang, D.S.: A study on fire data augmentation from video/image using the similar-label and f-guessed method. *Computer Science and Information Systems* (00), 11–11 (2024)
9. Kleiman, R., Page, D.:  $Auc_{\mu}$ : a performance metric for multi-class machine learning models. In: *International Conference on Machine Learning*. pp. 3439–3447. PMLR (2019)
10. Krummenacher, G., Ong, C.S., Koller, S., Kobayashi, S., Buhmann, J.M.: Wheel defect detection with machine learning. *IEEE Transactions on Intelligent Transportation Systems* 19(4), 1176–1187 (2017)
11. Kumar, M.P., Ashok, D.: A multi-level colour thresholding based segmentation approach for improved identification of the defective region in leather surfaces. *Engineering Journal* 24(2), 101–108 (2020)
12. Lee, K.H., Lee, H.W., Yun, G.J.: A defect detection framework using three-dimensional convolutional neural network (3d-cnn) with in-situ monitoring data in laser powder bed fusion process. *Optics & Laser Technology* 165, 109571 (2023)
13. Liu, J., Malekzadeh, M., Mirian, N., Song, T.A., Liu, C., Dutta, J.: Artificial intelligence-based image enhancement in pet imaging: Noise reduction and resolution enhancement. *PET clinics* 16(4), 553–576 (2021)
14. Lu, T., Wang, Z., Shen, Y., Shao, X., Tang, Y.: Defvae: A defect detection method for catenary devices based on variational autoencoder. *IEEE Transactions on Instrumentation and Measurement* (2023)
15. Ma, B., Zhu, W., Wang, Y., Wu, H., Yang, Y., Fan, H., Xu, H.: The defect detection of personalized print based on template matching. In: *2017 IEEE International Conference on Unmanned Systems (ICUS)*. pp. 266–271. IEEE (2017)
16. Mezher, A.M., Marble, A.E.: A novel strategy for improving robustness in computer vision manufacturing defect detection. *arXiv preprint arXiv:2305.09407* (2023)
17. Mezher, A.M., Marble, A.E.: Computer vision defect detection on unseen backgrounds for manufacturing inspection. *Expert Systems with Applications* 243, 122749 (2024)
18. Nagaraju, M., Chawla, P., Kumar, N.: Performance improvement of deep learning models using image augmentation techniques. *Multimedia Tools and Applications* 81(7), 9177–9200 (2022)
19. Ni, X., Liu, H., Ma, Z., Wang, C., Liu, J.: Detection for rail surface defects via partitioned edge feature. *IEEE Transactions on Intelligent Transportation Systems* 23(6), 5806–5822 (2021)
20. Ren, Z., Fang, F., Yan, N., Wu, Y.: State of the art in defect detection based on machine vision. *International Journal of Precision Engineering and Manufacturing-Green Technology* 9(2), 661–691 (2022)
21. Saberironaghi, A., Ren, J., El-Gindy, M.: Defect detection methods for industrial products using deep learning techniques: A review. *Algorithms* 16(2), 95 (2023)
22. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 4510–4520 (2018)
23. Sheng, T., Feng, C., Zhuo, S., Zhang, X., Shen, L., Aleksic, M.: A quantization-friendly separable convolution for mobilenets. In: *2018 1st Workshop on Energy Efficient Machine Learning and Cognitive Computing for Embedded Applications (EMC2)*. pp. 14–18. IEEE (2018)
24. Shipway, N., Barden, T., Huthwaite, P., Lowe, M.: Automated defect detection for fluorescent penetrant inspection using random forest. *NDT & E International* 101, 113–123 (2019)
25. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *Journal of big data* 6(1), 1–48 (2019)
26. Vrbačič, G., Pečnik, Š., Podgorelec, V.: Hyper-parameter optimization of convolutional neural networks for classifying covid-19 x-ray images. *Computer Science and Information Systems* 19(1), 327–352 (2022)

27. Wang, H., Wang, Y.: Improved glove defect detection algorithm based on yolov5 framework. In: 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). pp. 1192–1197. IEEE (2022)
28. Xu, B., Zhou, D., Li, W.: Image enhancement algorithm based on gan neural network. IEEE Access 10, 36766–36777 (2022)
29. Xu, C., Li, W., Cui, X., Wang, Z., Zheng, F., Zhang, X., Chen, B.: Scarcity-gan: Scarce data augmentation for defect detection via generative adversarial nets. Neurocomputing 566, 127061 (2024)
30. Zhang, K.: Using deep learning to automatic inspection system of printed circuit board in manufacturing industry under the internet of things. Computer Science and Information Systems 20(2), 723–741 (2023)
31. Zhang, Z., Sabuncu, M.: Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in neural information processing systems 31 (2018)
32. Zhou, X., Wang, Y., Xiao, C., Zhu, Q., Lu, X., Zhang, H., Ge, J., Zhao, H.: Automated visual inspection of glass bottle bottom with saliency detection and template matching. IEEE Transactions on Instrumentation and Measurement 68(11), 4253–4267 (2019)

**Yong Ren** is an Associate Professor with the School of Engineering, Applied Technology College of Soochow University, SuZhou, China. He received his Master's degree in Software Engineering from Soochow University. His research interests include artificial intelligence, machine vision, and image processing.

**Dong Liu** is a Bachelor of Computer Science from the Engineering School, Applied Technology College of Soochow University. His research interests include artificial intelligence, and machine vision.

**Sanhong Gu** is a Senior Engineer of Suzhou Dechuang Measurement and Control Technology Co., Ltd. His research interests include machine vision, and machine design.

*Received: September 11, 2024; Accepted: November 29, 2024.*

