

Article

Image Enhancement of Steel Plate Defects Based on Generative Adversarial Networks

Zhideng Jie ¹, Hong Zhang ^{2,*}, Kaixuan Li ¹, Xiao Xie ¹ and Aopu Shi ¹

¹ School of Mechanical Engineering, Jiangsu University of Technology, Changzhou 213001, China; 2022655018@smail.jsut.edu.cn (Z.J.); 2022655017@smail.jsut.edu.cn (K.L.); 2022655012@smail.jsut.edu.cn (X.X.); 2022655024@smail.jsut.edu.cn (A.S.)

² School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, China

* Correspondence: hazh0216@jsut.edu.cn; Tel.: +86-181-5173-7092

Abstract: In this study, the problem of a limited number of data samples, which affects the detection accuracy, arises for the image classification task of steel plate surface defects under conditions of small sample sizes. A data enhancement method based on generative adversarial networks is proposed. The method introduces a two-way attention mechanism, which is specifically designed to improve the model's ability to identify weak defects and optimize the model structure of the network discriminator, which augments the model's capacity to perceive the overall details of the image and effectively improves the intricacy and authenticity of the generated images. By enhancing the two original datasets, the experimental results show that the proposed method improves the average accuracy by 8.5% across the four convolutional classification models. The results demonstrate the superior detection accuracy of the proposed method, improving the classification of steel plate surface defects.

Keywords: image processing; dual-path attention mechanism; data augmentation; generative adversarial networks



Citation: Jie, Z.; Zhang, H.; Li, K.; Xie, X.; Shi, A. Image Enhancement of Steel Plate Defects Based on Generative Adversarial Networks. *Electronics* **2024**, *13*, 2013. <https://doi.org/10.3390/electronics13112013>

Academic Editors: Silvia Liberata Ullo and Li Zhang

Received: 16 April 2024

Revised: 12 May 2024

Accepted: 15 May 2024

Published: 22 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the progress of industrial technology, steel products have become widely utilized across numerous fields, such as aerospace engineering, machinery manufacturing, and the automobile industry. However, in the production process of steel plates, various external factors can lead to quality issues. Individual steel plate products often have quality problems, such as surface cracks, scratches, and inclusions. These problems greatly affect the corrosion resistance and wear resistance of steel plates, so the detection of surface defects on steel plates is particularly important [1–3].

To address these issues, researchers have developed various surface defect detection techniques. Previously, manual inspection was a primary method, but its limitations have become apparent with the evolution of industrial production and increasing quality demands. Manual inspections are labor-intensive, inefficient, and costly. Moreover, they are highly prone to human error and environmental factors, such as lighting changes, leading to inconsistent and unreliable results. These drawbacks make manual inspection unsuitable for modern, large-scale, complex industrial settings. In contrast, traditional computer vision employs industrial cameras to capture images, using standard image processing and machine learning methods for analysis. For instance, the study by Guo Hui and others introduced a defect feature extraction technique utilizing a Support Vector Machine (SVM) for steel plate defect identification [4]. Suvdaa et al. proposed an algorithm using the Scale-Invariant Feature Transform (SIFT) and voting strategy to solve this problem [5]. Although traditional computer vision methods have made progress in enhancing detection efficiency and accuracy, they still exhibit notable limitations. These methods typically rely on specific image features and algorithms, and they are highly sensitive to changes in steel plate

surface texture and lighting conditions, which limit their generalization and adaptability. Furthermore, these methods are confined to the detection of known defect types, and their performance is often unsatisfactory when confronted with new or unknown defect types [6]. In conclusion, the conventional approaches for identifying surface defects on steel plates, from manual examinations to classical machine vision methods, face inherent constraints in efficiency, precision, and adaptability. Such shortcomings not only impair inspection accuracy but also impede the quality assurance processes in manufacturing. Consequently, the development of a more proficient, exacting, and robust defect detection method is imperative to meet the rigorous standards of quality control required in today's manufacturing industries.

In contrast, deep learning methods automatically extract and learn the features of defect images through convolutional neural networks, which have the advantages of higher accuracy, speed, and the ability to adapt to different defect types [7]. For instance, scholars Zhang Guangshi et al. used the structure of a dense connected network (Dense Net) to improve the feature extraction ability [8]. Luo W and Zhang H proposed an efficient fine-grained image classification method that focuses on boosting the semantic strength of sub-features in global features. They achieved this through a clever set of regularization techniques of channel rearrangement and weighting combinations, and cleverly incorporated this mechanism into existing models without additional parameters, enabling end-to-end training. This approach has shown impressive results when relying only on image-level annotations [9]. Chen Y and Yang J aimed to enhance the robustness of the model against various noises and its ability to deal with occluded or corrupted images by introducing an adaptive noise dictionary. This research method focuses on the IRRPCA method to build an effective noise dictionary through which to improve the recognition rate of facial images [10]. In addition, although deep learning-based methods offer higher accuracy, speed, and adaptability to various defect types, there are still challenges in data enhancement processing for small-sample datasets, and the lack of diversity in training samples may lead to an insufficient generalization ability of the model. Therefore, it is necessary to propose a new method to overcome these limitations and further promote the development of steel plate surface defect detection technology.

The development of deep learning technology is based on the successive development of the Deep Boltzmann Machine (DBM), Pixel Recurrent Neural Network (PixelRNN), and generative adversarial network (GAN). Data enhancement using methods such as generative adversarial networks (GANs) on small-sample data has become a hot research topic [11]. For example, Goodfellow and others first proposed generative adversarial neural networks and applied them to handwritten digits and facial image generation [12]. Deep convolutional generative adversarial networks (DCGANs), proposed by Radford et al., are an improvement upon generative adversarial networks (GANs) [13] which significantly improve the generative ability and stability of GANs and offer a powerful tool for the study of image generation and related fields. Kun Liu, Aimei Li et al. proposed the combination of generative adversarial networks (GANs) with the One-Class Classifier and the utilization of GANs for automatic feature learning. The One-Class Classifier was used to detect abnormal samples, thus improving the accuracy and efficiency of surface defect detection of steel plates [14]. Shengqi Guan, Jiang Chang, et al. proposed an improved generative adversarial network (GAN) and EfficientNet to classify strip steel defects; they achieved efficient defect detection in the industrial field through automatic feature extraction and deep learning algorithms, yielding remarkable results [15]. Although these methods have improved the speed and accuracy of detection, they still face challenges when dealing with small-sample datasets, particularly as the model's generalizability may be limited when there is insufficient diversity in the training samples. To address this issue, this paper proposes a generative adversarial network incorporating a dual-path attention module and a U-Net structure. This paper aims to overcome the limitations of existing methods. In our study on generative adversarial networks (GANs), we have introduced a dual-path attention mechanism into the discriminator, along with the integration of the U-Net architecture

to enhance its performance. This improvement offers more accurate spatial consistency feedback to the generative model, significantly enhancing the clarity and detail of the generated images. Additionally, by fully leveraging the advanced learning capabilities of GANs, we address the challenges associated with data augmentation in small-sample datasets, thereby increasing the diversity of the training data and reducing the risk of overfitting. This enables the model to better adapt to unseen data, subsequently improving product quality and production efficiency, thus effectively propelling the advancement of metal surface defect detection technology [16].

2. Research Methods

In this study, we present a generative adversarial network (GAN) that is improved with a dual-path attention mechanism (DPAT) and incorporates a U-Net-based discriminator. This network, combined with a predictive model, effectively classifies defects, creating a system for generating and analyzing steel surface defect images. As illustrated in Figure 1, the integration of U-Net and the DPAT into the GAN improves the precision and quality of the images produced.

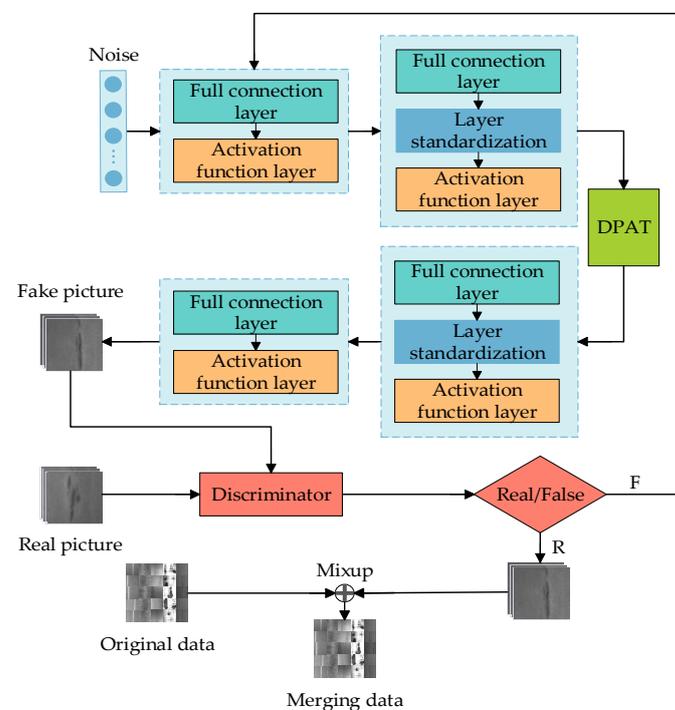


Figure 1. Generative model framework flowchart.

The generator features layers of fully connected units and activation functions, with a standardized integration layer for stable training. Importantly, embedding the DPAT within the generator's second layer boosts its capacity to pinpoint fine image details, refining the detection of minuscule defects. This prevents significant data loss in the output and accelerates the learning of data feature relationships.

We have updated the discriminator from a linear to a U-Net structure, enhancing its global and local decision-making capabilities. It accurately assesses image authenticity, performs detailed pixel-level classification, and offers continuous spatial feedback to the generator. This new discriminator improves the precision of the model, encouraging the generation of more intricate details in images and elevating the overall quality of the generated pieces.

In short, this structure's design balances detail with overall image quality and establishes a strong link between generative and discriminative models. The enhanced discriminator and DPAT not only sharpen the detail captured but also boost the model's

strength in handling complex images. Furthermore, each layer’s deliberate function and interaction theoretically allow for the precise detection of minor flaws and complex textures, thus raising the image generation quality without losing essential features.

2.1. Dual-Path Attention Module

This paper introduces the DPAT module to enhance the detection and classification of minor defects on steel surfaces through deep learning models [17]. By integrating the DPAT module into the GAN network and utilizing the outputs of the second layer of the generator and discriminator as inputs to the DPAT module, the model can learn the detailed texture of the image. The DPAT module processes self-attentive features by integrating average and maximum pooling layers, capturing feature correlations, and reducing its reliance on external data. This boosts the model’s ability to detect minor defects with heightened sensitivity and precision. The DPAT’s design focuses on effectively identifying subtle, often-missed defects without losing important image details. Its rapid internal correlation detection also improves the model’s grasp and application of specific image features, which are depicted in Figure 2.

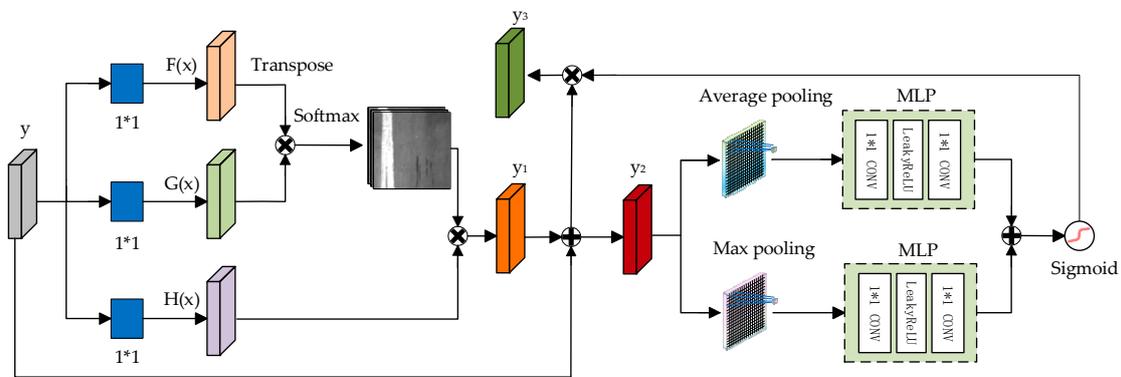


Figure 2. Dual-path attention module.

Firstly, the feature map y is obtained through feature extraction from the input defective image. Subsequently, a 1×1 convolution operation is applied to reduce the number of channels in the feature map y , resulting in a new feature map y_1 , and then y_1 is weighted and summed with the original y to obtain y_2 , as computed in Equation (1).

$$y_2 = V_{Att} \left\{ S \cdot M(x) \left[\frac{G(x)F(x)^T}{\sqrt{d_{H(x)}}} \right] \right\} + y \tag{1}$$

where S is the activation function, V_{Att} represents the attention computation, $F(x)$ and $G(x)$ are the feature vectors utilized for computing the attention weights, and $M(x)$ denotes the input feature vector. Subsequently, we employ mean pooling and maximum pooling techniques to compress the spatial dimension of the feature map y_2 , generating a one-dimensional vector. Subsequently, this one-dimensional vector is then forwarded to a multilayer perceptron (MLP), which consists of a 1×1 convolutional layer, a LeakyReLU activation function, and a 1×1 convolutional layer. Finally, the results derived from parallel processing through the average pooling layer and the maximum pooling layer are weighted and summed. Subsequently, the feature map is generated by applying the Sigmoid activation function and multiplying it with y_2 to obtain y_3 . The computed feature map y_3 is shown in Equation (2).

$$y_3 = Sig\{V_{MLP}[T(y)] + V_{MLP}[Q(y)]\} \times y_2 \tag{2}$$

In this process, V_{MLP} represents the result computed through the multilayer perceptron (MLP) layer and $T(y)$ represents the average pooled features, whereas $Q(y)$ represents the

maximum pooled features. Maximum pooling accentuates local details and helps to facilitate the capture of important local structures in the image, while average pooling preserves more global information and aids in understanding the overall scene. During the multiplication process, the parts with larger weights are emphasized and the parts with smaller weights are suppressed.

This mechanism enhances the performance of the model on a specific task by spatially aligning features at different locations, thereby enabling the module to concentrate more on areas that play a pivotal role in the task.

2.2. The Discriminative U-Net Module

Based on generative adversarial networks, the generator structure has been improved through the replacement of the original discriminative model structure with the U-Net structure [18]. This novel structure can handle both global and local information. This improvement facilitates the generative model in presenting richer image details and also augments the performance of the discriminative model. Furthermore, the architectural adjustments made to the discriminative model render it more resilient to spoofing attempts, and the incorporation of the U-Net structure aims to facilitate more information exchange and feedback to the generator, which makes the generator more effective in understanding the structure and content of the input image, ultimately leading to the generation of more accurate synthetic images that closely resemble real ones. The U-Net structure of the discriminator is shown in Figure 3.

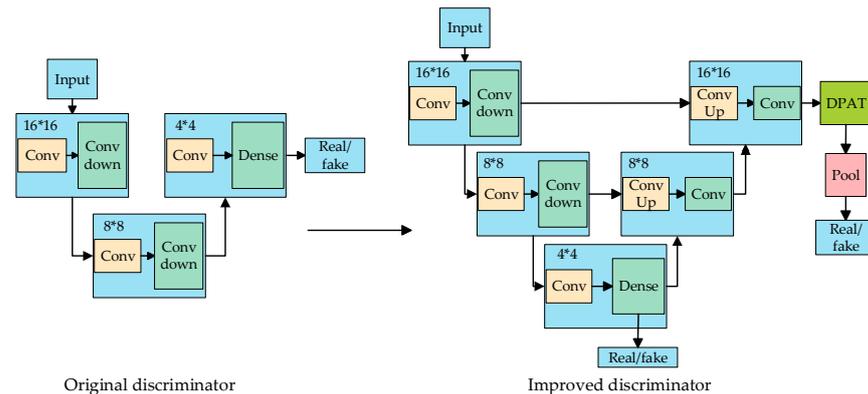


Figure 3. U-Net structure of discriminator.

Based on this, the loss function is adjusted by adopting the Cutout regularization method. The central concept of this method involves randomly selecting certain regions in the image and masking them (setting them to zero or assigning a designated fixed value). This approach encourages the model to place greater emphasis on the characteristics of the remaining visible regions, thereby enhancing the model’s generalization capability. The revised loss function of the discriminative model is shown in Equation (3).

$$L_D = L_{Denc} + L_{Ddec} \tag{3}$$

where L_{Denc} represents the loss function of the encoder portion of the discriminative model; it serves as a metric to quantify the dissimilarity between the discriminator’s output for the real data and that for the generated data. L_{Ddec} is the loss function of the decoder part of the discriminative model. This function assesses the discriminator’s capability to distinguish between the real and generated data, and it is elaborated on in detail in Equations (4) and (5).

$$L_{Denc} = -E_{x,z} \{ \ln D_{enc} [H(x)] + \ln \{ 1 + D_{enc} \{ H[G(z)] \} \} \} \tag{4}$$

$$L_{Ddec} = -E_{x,z} \{ \sum \ln \{ D_{dec} [H(x)] \} + \sum \ln \{ 1 - \{ D_{dec} \{ H[G(z)] \} \} \} \} \tag{5}$$

where D_{enc} and D_{dec} represent the encoder part and the decoder component of the discriminative model, respectively; H stands for the Cutout regularization method. The revised loss function of the generative model is presented in Equation (6).

$$L_G = -E_Z \{ \ln D_{enc} \{ H[G(z)] \} \} + \sum \ln \{ D_{dec} \{ H[G(z)] \} \} \quad (6)$$

3. Analysis and Comparison of Experimental Results

In our studies, we used the NEU Surface Defect Database and GC10-DET datasets for testing. The NEU dataset [19], created by Northeastern University, contains 1800 gray-scale images of steel defects like rolled oxide, plaque, and cracks, each 200×200 pixels. The GC10-DET [20], with 2279 images, captures real-world industrial defects of 10 varieties, including punched holes and creases, in a 2048×1000 resolution. These datasets challenge our defect detection model with varied defect types and sizes.

We have selected DCGAN [12], Style-GAN [21], SRGAN [22], WGAN [23], and the method proposed in this paper for comparison in our study. For each of the 16 different categories, 200 epochs of iterations were performed to generate 600 defective images respectively, and in the image quality assessment experiments, 10 evaluations were performed to select the optimal result.

3.1. Verification of Image Generation Quality

To verify the quality of the generated images, we conducted experiments using various algorithmic models to perform Fréchet Inception Distance (FID) tests on both the original dataset and the expanded dataset, respectively [24]. Test comparisons show the model's FID performance on generated defect images in Tables 1 and 2 below. A smaller FID value represents a higher quality of the generated image. The desired value (DV) represents the benchmark for the optimal FID scenario for the generated images. It signifies the Fréchet Inception Distance (FID) achieved when comparing the original dataset with itself.

Table 1. Comparison results of FID parameter metrics for images generated from NEU dataset.

Category	CR	IN	PA	PS	RS	SC
DCGAN	229.7	172.9	190.5	203.6	257.8	323.8
Style-GAN	96.4	150.3	123.1	93.8	112.2	146.7
OUR	76.3	116.2	90.0	75.9	98.7	104.8
DV	−2.5	−3.6	−4.6	−5.9	−2.9	−6.0

Table 2. Comparison results of FID parameter metrics for images generated from GC10-DET dataset.

Category	Ph	Wl	Cg	Ws	Os	Ss	In	Rp	Cr	Wf
WGAN	206.4	195.4	219.4	187.9	199.2	262.6	239.4	231.5	224.3	188.1
SRGAN	177.6	164.3	159.7	144.7	163.5	135.8	168.9	210.6	193.4	129.4
OUR	113.5	126.1	72.8	126.8	88.6	145.4	120.3	198.3	182.1	132.4
DV	−3.7	−4.9	−4.1	−5.4	−6.2	−3.2	−5.2	−7.7	−6.5	−4.7

When observing Table 1, we can see that our algorithm achieves a significant improvement under all six types of labels in the comparison results of the images generated from the NEU dataset. The reduction in FID values ranges from 10% to 20% relative to DCGAN and Style-GAN. Among them, the improvement is most evident under CR labeling, where our algorithm achieves an FID value of 76.3, compared to the values of 229.7 and 96.4 for DCGAN and Style-GAN, respectively. Furthermore, our algorithm achieves a greater improvement of 29% under SC labeling, followed by significant improvements under all other labels; for the indicators of IN, PA, PS, and RS, the OUR model improved by 32%, 53%, 63%, and 61% compared with the DCGAN model, respectively.

As can be seen from Table 2, after evaluating the performance of the three models WGAN, SRGAN, and OUR in terms of FID values, it is observed that the OUR model has

the best overall performance in all categories of defective image generation, with relatively low FID values, especially in the categories of Cg and Os, which demonstrate better defect identification and generation capabilities, but it performs poorly in the categories of Rp, Cr, and Ss. The reason for this is explained as follows: it is due to category imbalance and insufficient feature extraction. The SRGAN model also outperforms WGAN in most categories, especially in the In and Wf categories, which can better preserve image details and textures. In contrast, the WGAN model has relatively high FID values in most defect categories, generates lower-quality images, and performs poorly in the Ss, Wf, and Rp categories, which may be limited by its defect recognition and generation capabilities. In the indicators of Ph, Wl, and Ws, compared with the basic model, there is a 30%~40% improvement. Compared to the DV values in the ideal case, the OUR model is closer to the ideal FID values, followed by SRGAN, while WGAN deviates significantly from the ideal FID values.

Figures 4 and 5 display the results of a comparative analysis of images generated by multiple models.

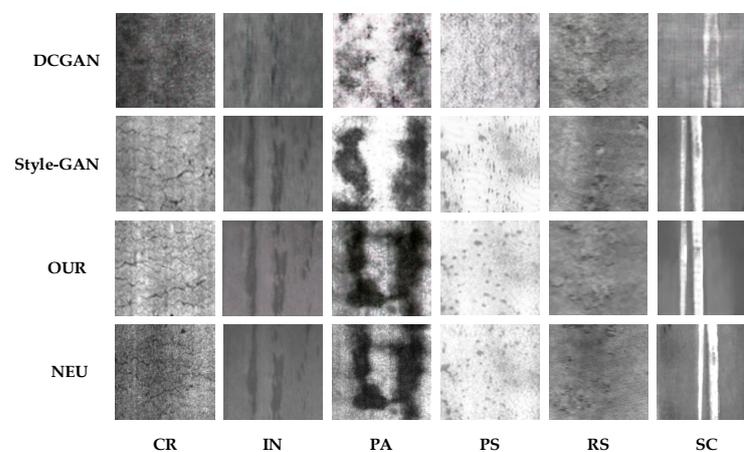


Figure 4. A comparison of the generated results of the model in the NEU dataset.

Analyzing the images generated in Figures 3 and 4, the results show that after 200 epochs of training, the OUR model's results are overall better, and the OUR-type images are intuitively closer to the original dataset, which suggests that the model performs well in capturing the core features and details of the dataset. It also exhibits greater similarity to the original dataset in terms of texture and local features. However, the features are not prominent on Cr and Ss types, which can be attributed to insufficient training samples, thereby resulting in poor model performance because of overfitting.

In contrast, the Style-GAN model has made significant progress in terms of overall performance, specifically in visual fidelity and diversity. Nevertheless, there remain certain deficiencies in the micro-level refinement of image details. In particular, regarding boundary clarity and micro-fine feature expression, the images generated by Style-GAN may exhibit edge blurring. This limitation may arise from the model's mechanism not being fully mature in efficiently capturing and accurately reconstructing local edge features during the iterative construction of the image. Observing the DCGAN-generated images, it is evident that DCGAN enhances the spatial structure generation of the images through the utilization of a deep convolutional structure. However, in the actual output, the general contour of the object often exhibits softer or even blurred characteristics, and the expressiveness of the key features is relatively weak. This results in a certain limitation in the recognizability of the generated images. The root cause of this limitation lies in DCGAN being insufficient at learning the deep and high-resolution features of the underlying dataset during the training stage, or because of the inefficient information transfer and retention in the generation path.

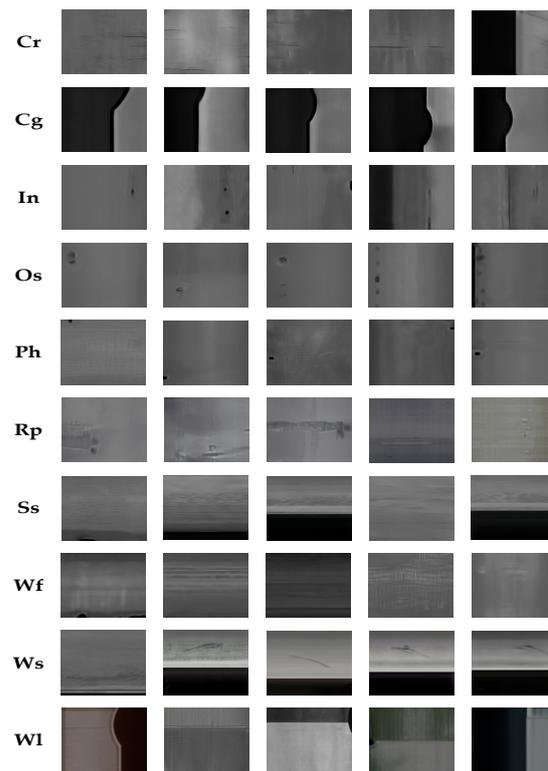


Figure 5. The OUR model's generated images in the GC10-DET dataset.

The experimental results show that the OUR model excels at fixing various data defects and enhancing diversity. It produces images that both complement and enrich the original dataset's content and distribution. However, its performance drops in complex situations, especially with rare sample categories. Despite this, the OUR model still offers significant improvement and potential for further development compared to the existing models.

3.2. Comparison of Model Prediction Results

To verify the performance of our dataset on the classification models, we utilized four experimentally commonly used models, ResNet, MobileNet, Vgg, and Transformer, for testing. The aim was to assess the improvement in prediction accuracy of the models before and after data augmentation [25]. The comparison results of the NEU dataset, both before and after augmentation, through several comparison tests, are presented in Table 3.

Table 3. Comparison of dataset before and after augmentation.

Metrics	ResNet		MobelNet-V3		Vgg		Transformer	
	Before	Improved	Before	Improved	Before	Improved	Before	Improved
F1-Score	92.92	96.77	95.64	96.77	94.08	96.89	82.39	85.54
Precision	93.95	96.54	95.65	97.23	94.33	96.91	84.23	86.31
Recall	93.00	95.78	95.73	96.66	94.11	96.89	81.69	83.61

From Table 3, it can be observed that the key performance indicators of F1-score, precision, and recall demonstrate significant improvements. The models achieved an average accuracy of over 85% for both F1-score and precision, whereas recall exhibited a positive growth trend of 1% to 2% when compared to the initial state. In Figures 6 and 7, it is evident that the augmented training set exhibits a smoother and more gradual performance growth compared to the original dataset throughout the model training process. The average increase in accuracy (ACC) across different models surpasses 2.7%, and the performance on the test set remains satisfactory. Upon analyzing the confusion matrix, it becomes intuitive that most categories are predicted with high accuracy; however, there exist certain

categories that are challenging to categorize and are thus misclassified to a certain degree. Overall, the improvement measures implemented in this study demonstrate substantial progress across various metrics, thereby offering a referential direction for optimizing and addressing specific category identification challenges.

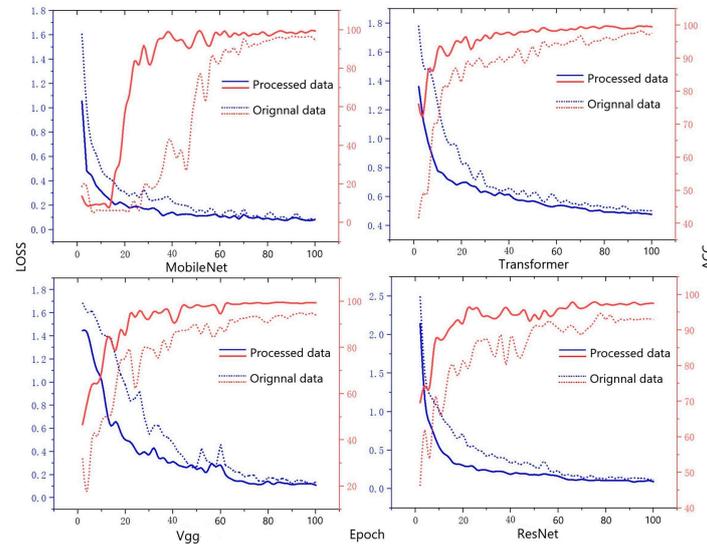


Figure 6. Predictive model training curve.

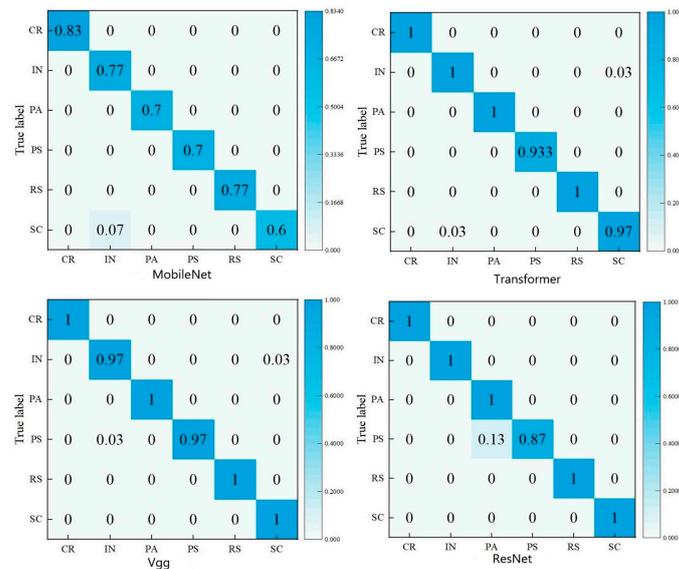


Figure 7. Test set confusion matrix.

Within the intricate architectures of deep learning models, a profound comprehension of the individual contribution of each component or parameter to the overall model performance is absolutely crucial. Such clarity assists in purging superfluous elements, fine-tuning the functionality of the model, and fostering a deeper intellectual engagement with how we can build upon the existing framework to achieve enhancement. The methodology encapsulated in this paper integrates multiple innovative modules and technological advances. In an effort to demystify the specific contributions of each segment to the ultimate performance, we have undertaken comprehensive ablation studies. We conducted controlled experiments with four model setups: the original model, the original model with DPAT data enhancement (Original + DPAT), the original model combined with the U-Net structure (Original + U-Net), and a model integrating both the DPAT and U-Net (Original +

DPAT + U-Net). We rigorously maintained control by training and testing all models on the same datasets to ensure fairness and comparability. During training, we closely monitored the key metric, mAP, to precisely assess each model’s performance trends.

As shown in Figures 8 and 9, the experimental findings reveal that the combined model incorporating both the DPAT and U-Net components (Original + DPAT + U-Net) outperforms all other experimental models in terms of recognition efficacy; the average accuracy (mAP) of the two sets is approximately 85% and 80%, respectively. mAP values improved by about 8.5% compared to the control group that relied only on the basic model. Of particular note is the fact that while models employing either a single DPAT enhancement strategy or the standalone U-Net architecture do exhibit some level of performance improvement over the original model, such enhancements are relatively less substantial when juxtaposed with those achieved by the composite model.

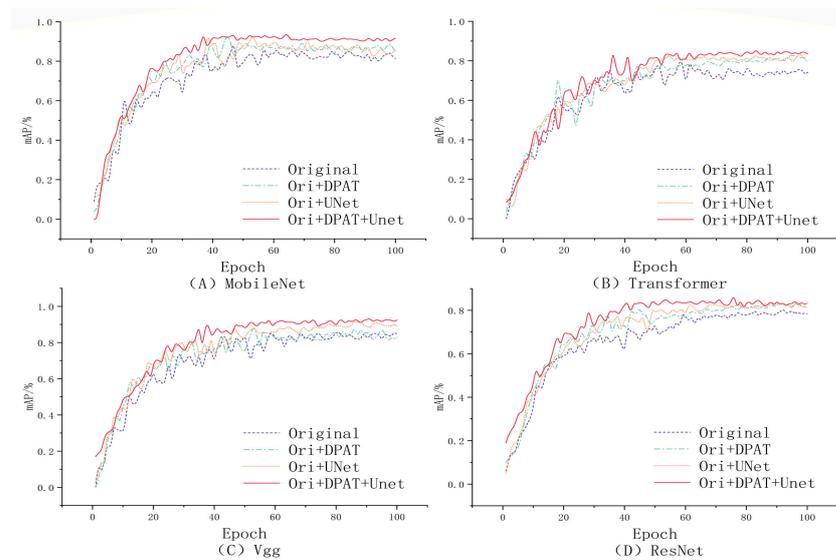


Figure 8. NEU dataset’s mAP values obtained for different models.

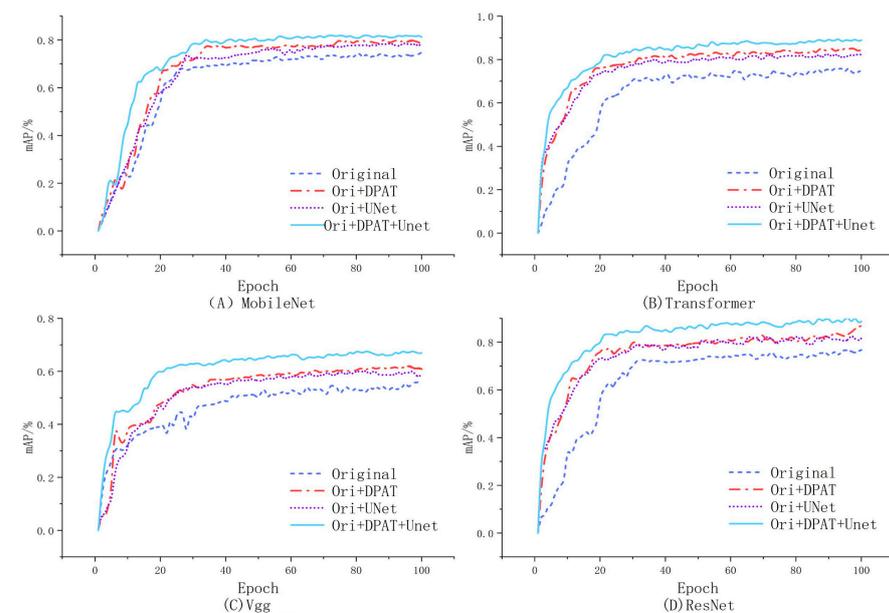


Figure 9. GC10-DET dataset’s mAP values obtained for different models.

Upon examining the Original + DPAT + U-Net model, we found that its performance is erratic in the early training stages, indicating a struggle to adapt initially. This could

be because of the model's heightened complexity from added data augmentation and the intricate U-Net design. As a result, it may need more iterations to fine-tune its parameters and learn detailed features. Moreover, using the DPAT and U-Net together could slow down convergence, especially with less-than-ideal initial parameter settings. Finally, the choice of training strategy also affects how swiftly and steadily the model adapts, with factors like improper learning rates or too much regularization possibly impeding effective training and stable progress.

In a further analysis, it is imperative to note the crucial impact that image resolution has on the difficulty of model training and the predictive results. While high-resolution images carry abundant information and subtle details, they also demand a higher learning and processing capacity from the model. In particular, in the GC10-DET dataset, the number of pixels in a single image increase by approximately 50 times compared to the NEU dataset. The prediction accuracy of the former is significantly lower than that of the latter. The amount of information that the model needs to process and learn significantly increases, causing a complexity surge that even data augmentation techniques struggle to fully compensate for, resulting in performance degeneration because of insufficient detail capture. Therefore, during the process of data augmentation and model training, the judicious selection and adjustment of image resolution, as well as the design of model architectures that can adapt to different resolution requirements, become key factors in enhancing model performance.

The ablation studies conducted have further substantiated the pivotal role of the bidirectional attention mechanism and the U-Net architecture in determining the results, as well as the robustness of our model when addressing the challenges of learning from small sample sizes. These experiments furnish our approach with decisive quantitative backing and carve out a valuable perspective for future research endeavors. We hope that through these ablation studies, researchers will acquire a clear appreciation of why each design step was imperative and how they collectively augment the performance of our model. The transparency and rigor of our experimental approach are intended to serve as a reference point for researchers who may wish to replicate or refine our methods in future studies.

4. Conclusions

In this paper, the problem of a limited number of data samples, which affects the accuracy of detection, persists in classifying steel plate surface defect images under small sample conditions. A data enhancement method based on the generative adversarial network is proposed. By introducing a dual-path attention module mechanism and improving the discriminator model, the target detection ability in this scenario is effectively improved. The experimental results demonstrate that the algorithm proposed in this paper outperforms the comparison algorithm, effectively enhancing the detection accuracy while guaranteeing the detection efficiency, and improving the accuracy of defect detection on the steel surface under small sample conditions. The data enhancement method proposed in this paper has a positive impact on the overall predictive performance of the model and is expected to provide strong technical support for the development of intelligent defect detection in the steel manufacturing industry.

Author Contributions: Methodology, Z.J.; validation, Z.J.; writing—original draft, Z.J.; formal analysis, H.Z.; investigation, K.L.; data curation, X.X.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Postgraduate Research & Practice Innovation Program of Jiangsu Province; Project Number: XSJCX23_03.

Data Availability Statement: Some or all of the data and models generated or used during this study are available in a repository or online.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Huang, B.; Liu, J.; Liu, X.; Liu, K.; Liao, X.; Li, K.; Wang, J. Improved YOLOv5 Network for Steel Surface Defect Detection. *Metals* **2023**, *13*, 1439. [[CrossRef](#)]
2. Dou, Z.; Hu, C.; Li, Q.; Zheng, L. Improved YOLOv7 Algorithm for Small Sample Steel Plate Surface Defect Detection. *Comput. Eng. Appl.* **2023**, *59*, 283–292.
3. Xie, Q.; Zhou, W.; Tan, H.; Wang, X. Surface Defect Recognition in Steel Plates Based on Improved Faster R-CNN. In Proceedings of the 2022 41st Chinese Control Conference (CCC), Hefei, China, 25–27 July 2022; pp. 6759–6764.
4. Guo, H.; Xu, W.; Liu, Y. Surface Defect Detection of Steel Plates Based on Support Vector Machine. *J. Donghua Univ. (Nat. Sci. Ed.)* **2018**, *44*, 635–639.
5. Suvdaa, B.; Ahn, J.; Ko, J. Steel surface defects detection and classification using SIFT and voting strategy. *Int. J. Softw. Eng. Its Appl.* **2012**, *6*, 161–166.
6. He, C.; Zhang, K.; Yu, R. Research review on the application of machine vision in Rail surface disease detection. *New Mater. High-Speed Railw.* **2024**, *3*, 7–13.
7. Wu, G.; Ning, X.; Hou, L.; He, F.; Zhang, H.; Shankar, A. Three-dimensional softmax mechanism guided bidirectional GRU networks for hyperspectral remote sensing image classification. *Signal Process.* **2023**, *212*, 109151. [[CrossRef](#)]
8. Zhang, G.S.; Ge, G.Y.; Zhu, R.H.; Sun, Q. Gear defect detection based on improved YOLOv3 network. *Laser Optoelectron. Prog.* **2020**, *57*, 153–161.
9. Luo, W.; Zhang, H.; Li, J.; Wei, X.S. Learning semantically enhanced feature for fine-grained image classification. *IEEE Signal Process. Lett.* **2020**, *27*, 1545–1549. [[CrossRef](#)]
10. Chen, Y.; Yang, J.; Luo, L.; Zhang, H.; Qian, J.; Tai, Y.; Zhang, J. Adaptive noise dictionary construction via IRRPCA for face recognition. *Pattern Recognit.* **2016**, *59*, 26–41. [[CrossRef](#)]
11. Zhou, F.; Jin, L.; Dong, J. Review of convolutional neural networks. *J. Comput. Sci.* **2017**, *40*, 1229–1251.
12. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]
13. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv* **2015**, arXiv:1511.06434.
14. Liu, K.; Li, A.; Wen, X. Steel surface defect detection using GAN and one-class classifier. In Proceedings of the IEEE 2019 25th International Conference on Automation and Computing (ICAC), Lancaster, UK, 5–7 September 2019; pp. 1–6.
15. Guan, S.; Chang, J.; Shi, H. Strip steel defect classification using the improved GAN and EfficientNet. *Appl. Artif. Intell.* **2021**, *35*, 1887–1904. [[CrossRef](#)]
16. Cheng, X.; Xie, L.; Zhu, J.; Hu, B.; Shi, S. Overview of Generative Adversarial Networks (GANs). *Comput. Sci.* **2019**, *46*, 74–81.
17. Zhou, Y.; Yuan, Z.; Chen, H. Solar cell defect generation based on meta-learning and dual path attention. *Acta Energetica Solaris Sin.* **2023**, *44*, 85–93.
18. Jiang, Y.; Zhu, B. Emote sensing image data enhancement based on generative adversarial networks under small sample conditions. *Laser Optoelectron. Prog.* **2021**, *58*, 0810022. [[CrossRef](#)]
19. Bao, Y.; Song, K.; Liu, J.; Wang, Y.; Yan, Y.; Yu, H.; Li, X. triplet-graph reasoning network for few-shot metal generic surface defect segmentation. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11.
20. Lv, X.; Duan, F.; Jiang, J.J.; Fu, X.; Gan, L. Deep metallic surface defect detection: The new benchmark and detection network. *Sensors* **2020**, *20*, 1562. [[CrossRef](#)] [[PubMed](#)]
21. Zhang, Y.; Niu, J.; Huang, Z.; Pan, C.; Xue, Y.; Tan, F. High-Precision Detection for Sandalwood Trees via Improved YOLOv5s and StyleGAN. *Agriculture* **2024**, *14*, 452. [[CrossRef](#)]
22. Yu, W.; Zhou, H.; Liu, Y.; Yang, Y.; Shen, Y. Super-Resolution Reconstruction of Particleboard Images Based on Improved SRGAN. *Forests* **2023**, *14*, 1842. [[CrossRef](#)]
23. Abdulraheem, A.; Suleiman, J.T.; Jung, I.Y. Generative Adversarial Network Models for Augmenting Digit and Character Datasets Embedded in Standard Markings on Ship Bodies. *Electronics* **2023**, *12*, 3668. [[CrossRef](#)]
24. Zhao, C.; Shuai, R.; Ma, L.; Liu, W.; Wu, M. Generation and Classification of Skin Cancer Images Based on Self-Attention StyleGAN. *J. Comput. Eng. Appl.* **2022**, *58*, 111.
25. Liu, K.; Wang, D.; Rong, M. X-ray image classification algorithm based on semi-supervised generation adversarial network. *Acta Opt. Sin.* **2019**, *39*, 0810003.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.