Research on Engineering Structures & Materials
www.jresm.org
Research Group

Research Article

Application of a GAN model incorporating transfer learning methods for the detection of forming defects in metallic materials

Gang Xu [*,1,a], Jiayu Xie [2,b], Chao Qiu [2,c]

[1]*Engineering Techniques Training Center, Civil Aviation Flight University of China, Guanghan, China*
[2]*College of Aviation Engineering, Civil Aviation Flight University of China, Guanghan, China*

Article Info

Article History:

Received 10 July 2026
Accepted 14 Aug 2026

Keywords:

Transfer learning;
Generative adversarial;
Metallic materials;
Forming defect detection;
Deep learning

Abstract

Defect detection is essential for quality assurance in metal forming, but traditional methods often suffer from process complexity and imbalanced datasets. This study proposes a GAN-based defect detection framework integrating transfer learning to improve accuracy and robustness. A pre-trained GAN is first employed to augment non-defective metal samples, enhancing data diversity, followed by transfer learning-based defect classification. Experimental results demonstrate that the improved auxiliary classifier GAN achieves an average classification accuracy of 95.98%, outperforming traditional auxiliary classifier GAN (91.08%) and ResNet-50 (88.65%). Moreover, the proposed model obtains a Mode Score of 1.386, higher than Res-Wasserstein GAN (1.374) and Res-ACGAN (1.361), indicating improved quality and diversity of generated samples. The model is further applied to defect detection of aluminum profiles and hot-rolled steel strips, where it effectively identifies and classifies various defect types with enhanced detection efficiency and reliability. These results demonstrate that the proposed AI-driven approach can improve intelligent quality control in metal processing, reduce potential safety risks caused by undetected defects, and provide an effective solution for advanced manufacturing applications.

1. Introduction

Detection of forming defects in metal materials (DFDMM) is essential to ensure product quality and productivity. However, conventional DFDMM methods, such as rule-based machine vision techniques, are limited by their adaptability to complex backgrounds or non-rule defects, robustness to external factors like lighting changes, and flexibility to quickly adapt to new defects or products. These limitations are further compounded by the challenges of uneven data sampling and complex shaping procedures in metal materials [1-2]. In recent years, novel techniques like Generative Adversarial Networks (GANs) and Transfer Learning (TL) have been introduced and present new avenues for DFDMM solutions. GANs are potent deep learning frameworks that generate lifelike data samples through adversarial training, which has shown promise in addressing the issue of limited training data [3-4]. TL, on the other hand, expedites the learning process and bolsters the model's performance by utilizing an already trained model for a new assignment, which can be particularly useful when dealing with small sample sizes. By combining these two techniques, it is possible to improve the performance of DFDMM [5-6]. This study explores how GANs integrated with TL address DFDMM challenges. Pre-trained GANs first augment defect-free metal samples to diversify training data, then TL applies the generator to DFDMM tasks. The novelty lies in handling multiple sub-tasks (classification and localization) jointly, unlike traditional methods that treat them separately. By sharing parameters, the GAN-TL model enables

*Corresponding author: 15883449712@163.com
[a]orcid.org/0009-0008-1972-6166; [b]orcid.org/0000-0002-4790-8265 ; [c]orcid.org/0000-0002-7490-0306
DOI: http://dx.doi.org/10.17515/resm2026-1850ai0710rs

knowledge transfer across tasks, boosting overall performance. This approach overcomes traditional limitations while leveraging GANs and TL strengths. In conclusion, GAN-TL models innovatively utilize TL, synthetic data generation, multi-task learning, and non-uniform data management to enhance performance and practicality, offering a robust solution for industrial defect detection.

2. Related Works

GANs have been widely used in data augmentation across various fields. For instance, Lei et al. [7] proposed a deep convolutional GAN for reconstructing lost data in structural health monitoring, demonstrating that the reconstructed signals closely matched the original signals in both time and frequency domains. Similarly, Cai et al. [8] utilized GANs to generate brain scans for deep learning models when certain MRI sequences were missing, showing that the generated images could effectively serve as inputs for these models. These studies highlight the potential of GANs in generating high-quality synthetic data, which was particularly useful in scenarios with limited data availability. However, these applications primarily focused on specific domains and did not directly address the challenges of metal material defect detection. In the context of metal materials, GANs have been used for generating synthetic defect images to augment training datasets. For example, Hu et al. [9] proposed a deep convolutional GAN for unsupervised fabric defect detection, which showed promising results in generating realistic defect samples. Despite these advancements, existing GAN-based methods often failed to effectively handle the complexities of metal material defect detection, such as the high variability in defect types and the need for accurate defect localization.

TL has been effectively employed in cross-domain detection tasks to leverage pre-trained models and adapt them to new domains with limited labeled data. Rajabi et al. [10] introduced a GAN-based model for generating tabular data, which demonstrated improved stability compared to traditional GANs. This approach was particularly relevant in scenarios where data from different domains exhibited significant distribution differences. Similarly, Krause et al. [11] used conditional GANs to synthesize histological images that retained genetic information, showcasing the potential of TL in medical imaging. In the context of defect detection, TL has been used to adapt models trained on one type of material to another. For instance, Baumgartl et al. [12] proposed a process monitoring architecture for metal parts manufactured using laser powder bed fusion, which leveraged TL to adapt models trained on one type of metal to another. However, these studies primarily focused on adapting models within the same domain and did not address the unique challenges of metal material defect detection, such as the need for simultaneous defect classification and localization.

Combining GANs and TL has shown promise in addressing the limitations of individual techniques. For example, Yang et al. [13] proposed a texture-conditional GAN for infrared and visible image fusion, which demonstrated superior performance in generating high-quality fused images. This approach leveraged the strengths of GANs in generating realistic data and TL in adapting models to new tasks. Similarly, Jabbar et al. [14] proposed a solution for stabilizing GAN training using unsupervised domain adaptation (DA), which showed improved stability and reduced mode collapse. In the context of metal material defect detection, combining GANs and TL could address the challenges of limited data and domain differences. However, existing studies often focused on individual tasks such as defect classification or localization, without fully leveraging the potential of multi-task learning. For instance, Jing et al. [15] proposed an efficient convolutional neural network for fabric defect detection, which achieved high segmentation accuracy and detection speed. Despite these advancements, existing methods often fail to fully exploit the correlation between multiple defect detection tasks, leading to suboptimal performance.

Despite progress in GANs and TL for defect detection, existing studies have limitations in handling metal material defect complexities [16-17]. They often focus on single tasks, ignore task correlations, and struggle with limited data and domain differences. To address this, the study proposes a GAN-TL approach for multi-task defect detection, leveraging GANs for synthetic data generation and TL for DA, while incorporating multi-task learning to improve performance. This maximizes task correlations, enhancing accuracy and robustness. Contributions: (1) pre-trained

GANs generate synthetic defect-free samples to diversify training. (2) TL adapts models to new tasks. (3) multi-task integration enables knowledge transfer. (4) It achieves superior accuracy and robustness under limited data and domain shifts. This provides a comprehensive solution for metal defect detection.

3. Defect Detection (D-D) Algorithm Based on GAN and TL

3.1. Improved Aided Classification GAN Defect Model

GAN, as a class of generative models that approximate the real data distribution, can attenuate the impact of small samples and class imbalance on the detection accuracy by means of data expansion. As depicted in Fig. 1, GAN is composed of a generator and a discriminator portion [18].

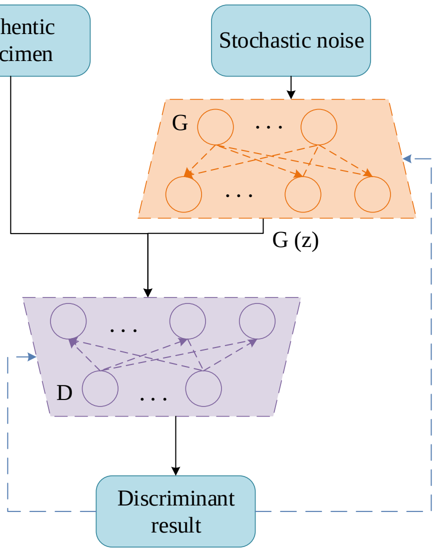


Fig. 1. The general framework of GAN

In Fig. 1, the noise z with P_z prior distribution law is fed into the generator to obtain a generated fake sample $G(z)$ and the real data x_r are used to discriminate the authenticity using the discriminator D, which ultimately optimizes the network by back propagation through the adversarial loss. The ultimate goal of the GAN is to make the generated of the sample distribution P_g as close as possible to the real sample P_r distribution coincidentally. In Equation (1), the GAN's overall objective function is displayed.

$$\min_G \max_D L(G,D) = E_{z\sim P_z}[log(-D(G(z))) + 1] + E_{x\sim P_r}[log\, D\,(x)] \tag{1}$$

In equation (1), G denotes the false sample and E denotes the expectation. Equation (2) illustrates the discriminator's objective function as well as the two networks' respective objective functions as the overall objective function.

$$\max_D L(D) = E_{z\sim P_z}[log(-D(G(z))) + 1] + E_{x\sim P_r}[log\, D\,(x)] \tag{2}$$

When the input is a produced sample, the discriminator minimizes the score $D(G(z))$, while D maximizes the discriminant score $D(x)$ in equation (2) when the input is a real sample. Equation (3) displays the generator's objective function.

$$\min_G L(G) = E_{z\sim P_z}[log(-D(G(z))) + 1] \tag{3}$$

Auxiliary Classifier GAN (ACGAN) is a class-conditional extension model of GANs. ACGAN adds a labelling constraint c to the generator network of GAN, which forms a joint hidden-layer representation with the input noise z, constraining the generator to generate samples of a specific class. Meanwhile, it provides different categories of control in the image generation process by adding classifiers to the discriminator, which enables both true and false discrimination of samples as well as the classification task. The adversarial loss representation is shown in equation (4).

$$L_{adv}(G,D) = E_{z\sim P_z, c_g\sim Pc_g}[log(-D(G(z,c_g))) + 1] + E_{x\sim P_r}[log\, D\,(x)] \tag{4}$$

In equation (4), Pc_g is the distribution of category labels of the generated samples and c_g are the category labels of the generated data. The final representation of the generator objective function of ACGAN is shown in equation (5).

$$L(G) = L_{adv}(G,D) + \lambda_r L_{cls}^{r}(C) + \lambda_g L_{cls}^{g}(C) \tag{5}$$

In equation (5), λ_r and λ_g are the hyperparameters of the loss of real and synthetic data classifiers. The discriminator objective function of ACGAN is shown in equation (6).

$$\underset{G}{min}\underset{D}{max}L(G,D) = E_{z\sim P_z}[log(-D(z)) + 1] + E_{x\sim P_r}[log\, D\,(x)] \tag{6}$$

The general framework diagram of ACGAN is shown in Fig. 2.

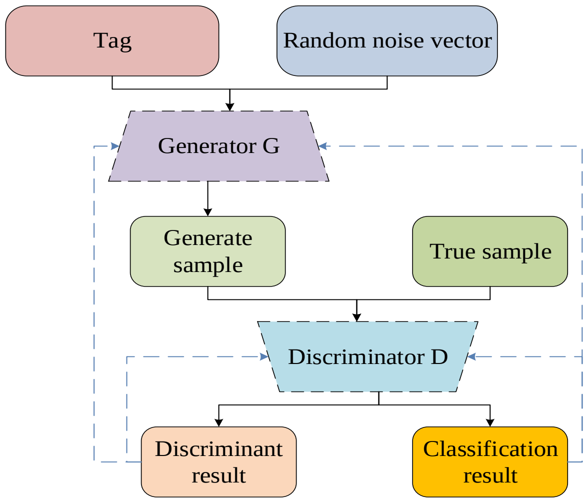


Fig. 2. Overall framework diagram of ACGAN

The improved ACGAN generator adopts an encoder-decoder structure. The encoder consists of 4 convolutional layers, with the convolution kernel sizes of each layer being 5×5, 5×5, 3×3, and 3×3 respectively, and the stride sizes being 2, 2, 1, and 1 respectively, and the channel numbers being 64, 128, 256, and 512 respectively. The output feature map sizes are 128×128, 64×64, 64×64, and 64×64 respectively. Each layer is followed by batch normalization and LeakyReLU activation. The decoder consists of 4 deconvolutional layers, which are symmetrical to the encoder. The convolution kernel sizes of each layer are 3×3, 3×3, 5×5, and 5×5 respectively, and the stride sizes are 1, 1, 2, and 2 respectively, and the channel numbers are 256, 128, 64, and 3 respectively. The output feature map sizes are 64×64, 64×64, 128×128, and 256×256 respectively. The last layer uses the Tanh activation function to output the generated image. The discriminator consists of 5 convolutional layers, with the convolution kernel size of each layer being 4×4, and the stride size being 2, and the channel numbers being 64, 128, 256, 512, and 1 respectively. The output feature map sizes are 128×128, 64×64, 32×32, 16×16, and 8×8 respectively. Each layer is followed by spectral norm normalization and LeakyReLU activation, and the last layer outputs the binary classification score. Residual connections are introduced after each convolutional layer of both the generator and the discriminator to alleviate the problem of gradient disappearance in deep networks and enhance the ability of feature reuse. Each layer is followed by spectral norm normalization and LeakyReLU activation. Both networks integrate the channel attention module, with the embedding position being after each convolutional layer. During training, the optimizer uses Adam. The learning rates of the generator and discriminator are both set to 0.0002, the batch size is set to 32, and the total training rounds are 200. After each training round, the performance of the current model is evaluated on the validation set, and the model parameters with the highest validation accuracy are saved for the final test. To stabilize the training, the update frequency of the discriminator is set to update the generator once after each update of the discriminator, and the gradient penalty coefficient is set to 10. The specific formula of gradient penalty is shown as equation (7).

$$\text{Gradient Penalty} = \lambda(\|\nabla_{\hat{x}}D(\hat{x})\|_2 - 1)^2 \tag{7}$$

In equation (7), $\hat{x}$ is the interpolation sample between the real sample and the generated sample, and λ is the weight parameter of the gradient penalty, which is usually set to 10. This gradient penalty mechanism ensures the stability of the training process and the diversity of the generated samples by limiting the gradient norm of the discriminator. The flowchart of the D-D algorithm for metal surfaces based on improved ACGAN is shown in Fig. 3.

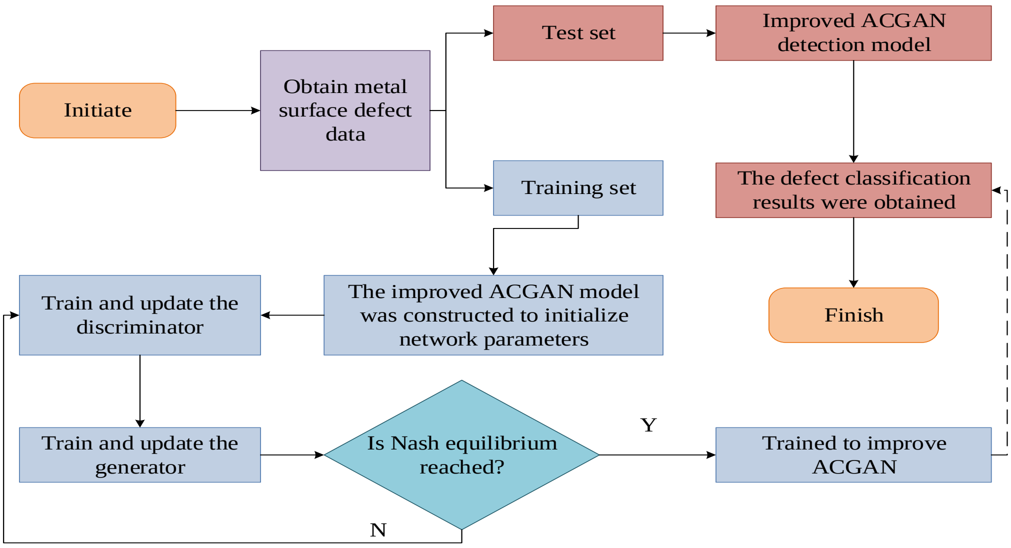


Fig. 3. Improved detection flow chart of ACGAN algorithm

In Fig. 3, a dataset of defect images of metal surfaces is acquired by an image acquisition device. Second, an attention mechanism is added to the discriminator and generator to improve the generation and learning of flawed features. To limit the gradient anomaly of the model and solve the pattern collapse problem of GAN, a penalty function for regret minimization is simultaneously added to the loss function of the discriminator. The improved ACGAN network is trained until the Nash equilibrium, the improved ACGAN model training is complete, the trained model and parameters are saved. For the illumination variations and imaging interference in industrial scenarios, the model integrates a multi-scale adaptive brightness equalization module in the preprocessing stage to correct the gray-scale distribution of the image. In the feature extraction stage, an angle-invariant feature extraction branch is set up to reduce the feature distortion caused by the deviation of the shooting angle. Moreover, a denoising convolution layer is adopted to suppress motion blur and electromagnetic noise.

3.2. Defect Classification Model Based on Adversarial Generation and Unsupervised DA

By leveraging source domain knowledge to guide target domain tasks, TL reduces reliance on labeled data, removes the need for independent and identically distributed training/testing data in traditional ML, and addresses domain migration and insufficient testing samples. DA, a TL model, tackles detection issues caused by dataset distribution differences. DA aims to build models with minimal or no labeled target data by aligning source-to-target distributions. Current DA methods are widely studied but perform poorly on defect data with large background distribution variations in real industrial settings [19]. To address this, the study combines GAN and unsupervised DA to propose AGAN for metal surface defect detection. As illustrated in Fig. 4, the proposed AGAN architecture primarily consists of three components: task categorization, sample creation, and feature extraction.

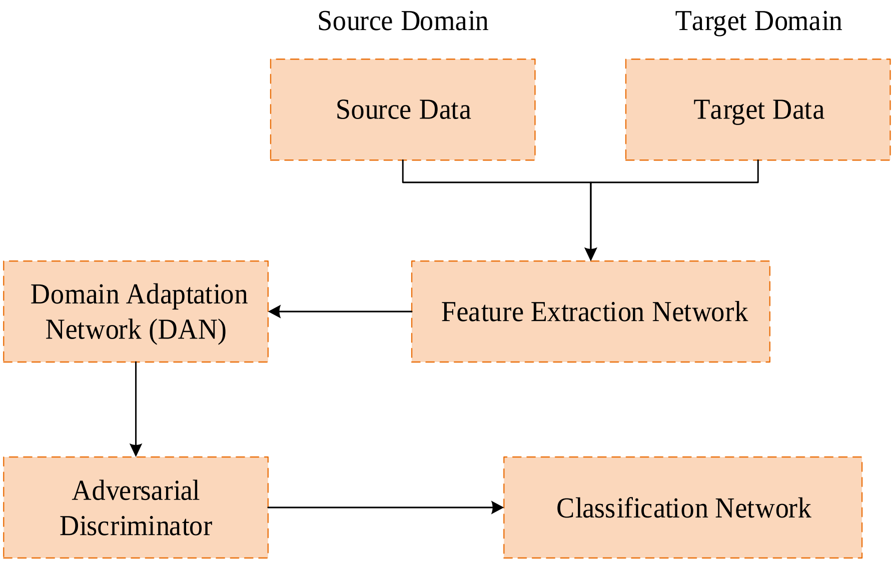


Fig. 4. Adaptive network model diagram against generation

During the TL based on AGAN, the shared encoder is initialized using the pre-trained parameters from the source domain dataset. During the training on the target domain, all convolutional layers of the shared encoder are fine-tuned with a lower learning rate to adapt to the specific features of the target domain while retaining the general feature representation. The domain classifier is trained from scratch because its task of distinguishing between the source domain and the target domain is specific to cross-domain scenarios. The category classifier is also trained from scratch, learning to map domain-invariant features to defect categories using only the source domain labels. The private encoder and decoder of the source domain are jointly fine-tuned with the shared encoder to maintain the reconstruction quality. No layers are frozen during the adaptation process because experiments have shown that the full fine-tuning of the shared encoder can achieve better domain alignment effects compared to freezing the shallow layers. The study chose to build a self-trained pre-training model on the source domain metal dataset instead of using publicly available pre-training weights. This decision is mainly based on the following considerations: Both the source domain and the target domain consist of images of surface defects of metal materials, and they share similarities in texture distribution and defect morphology. The pre-training parameters of the source domain can provide more effective feature initialization for the target domain. The natural image features of the public dataset differ significantly from the metal grayscale textures, and direct transfer has limited effectiveness and high computational costs, which is not conducive to lightweight deployment in industrial scenarios.

In the AGAN fine-tuning process, the learning rate of the shared encoder is set to 0.0001, the learning rates of the private encoder, decoder and both classifiers are set to 0.001, Adam is selected as the optimizer, the batch size is set to 32, and the total training rounds are 100. The shared encoder is updated once every two updates of the domain discriminator to maintain the stability of the adversarial training. All layers participate in the fine-tuning, no frozen layers are set, and the full fine-tuning strategy remains consistent under all experimental conditions. AGAN's feature enhancement method enhances the model's ability to identify and classify defects by integrating additional defect feature vectors, as shown in equation (8).

$$\tilde{f}_s = f_s + \lambda(L_{cls})(f_s^{-} - f_s) \quad (8)$$

In equation (8), f_s represents the feature vector of positive class samples, which is the feature representation corresponding to the target defect category. f_s^{-} represents the feature vector of negative class samples, which is the feature representation of non-defect or other non-target categories. $\tilde{f}_s$ represents the new feature vector generated based on the source domain features through the embedding enhancement method, L_{cls} represents the classification loss of the source domain data, and λ is a weight parameter factor used to balance the original classification loss and the enhanced feature classification loss. In this way, the model can not only utilize the existing feature vectors, but also dynamically integrate new feature vectors through the feedback during the training process, thereby improving the generalization ability and classification accuracy of the model. During the training process, the newly added feature vectors are sent to the classifier to predict their category labels. Subsequently, combined with the labels of these new features and the prediction results, the classification error is jointly calculated. Based on this classification error, the parameters responsible for feature extraction and classification in the encoder are adjusted. The training process of the specific feature enhancement module is shown in Fig. 5.

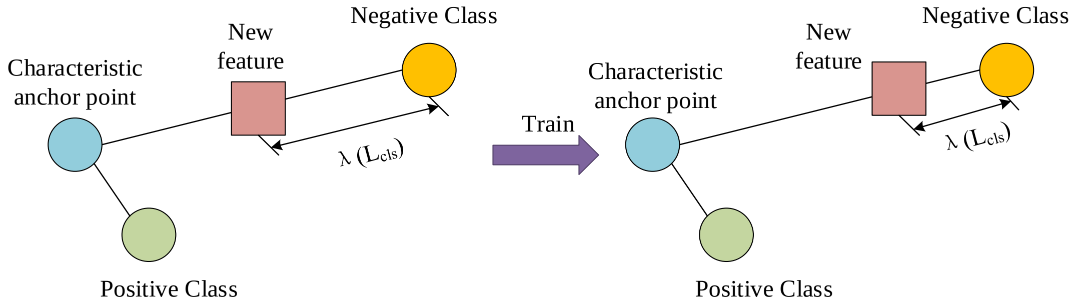


Fig. 5. Adaptive feature enhancement process

The new feature vectors are fed into the classifier to obtain the predicted labels, and the classification loss of the embedded samples is calculated jointly with the predicted labels by inputting labels for the new features, updating the parameters of the encoder's feature extraction part and classification part. The reconstruction loss of the network is determined using the mean square error MSE, and the final reconstruction loss is displayed in equation (9) in order to enhance the codec network's ability to reconstruct the source domain data.

$$L_{recon} = \frac{1}{n}\sum_{i=1}^{n}(x_s - x_s^{'})^2 + \frac{1}{n^2}\sum_{i=1}^{n}(x_s - x_s^{'})^2 \tag{9}$$

In equation (9), n is the sample size, and the private encoder network and decoder network are updated by the reconstruction loss. Then, the study promotes the decoder's ability to better restore the feature information of the source domain data while doing so. Equation (10), which is based on the adversarial network concept, displays the global domain alignment loss.

$$L_{global} = E_{x_t \sim P_t}[log\, D_{domain}\,(x_t)] + E_{x_S \sim P_S}[log(\,1 - D_{domain}(G(En_s(x_s^{'}))))] \tag{10}$$

In equation (10), E is the expectation, En_s is the shared encoder, and D_{domain} is the domain classifier. Based on the adversarial concept, the decoder aims to preserve the source domain's feature information in the reconstructed samples. This encourages the reconstructed data to retain features shared across domains. Therefore, the loss of the discriminator is shown in Equation (11).

$$L_D = -L_{global} \tag{11}$$

The loss of the decoder is shown in equation (12).

$$L_G = E_{x_S \sim P_S}[log(\,D_{domain}(G(En_s(x_s^{'}))))] \tag{12}$$

Equation (13) illustrates the classification loss, which includes the classification loss of the data from the source domain and the classification loss of the data produced by the source domain.

$$L_s = E_{x_S \sim P_S}[L_C(C_{cls}(En_g(x_s))|c = c_s)] + E_{x_S{}' \sim P_S{}'}[L_C(C_{cls}(En_g(x_s{}'))|c = c_s)] \tag{13}$$

In equation (13), $P_S{}'$ represents the distribution of the data for the reconstructed samples in the source domain, and c_s represents the labelling details for the data in the source domain. In addition to the above classification loss, there is also a classification loss consisting of new features embedded by the feature-based embedding enhancement method, as shown in equation (14).

$$L_{emb} = E_{\tilde{f}}[L_C(C_{cls}(\tilde{f})|c = \tilde{c})] \tag{14}$$

In equation (14), A denotes the labelling information of the newly embedded features. Therefore, the final classification loss is shown in equation (15).

$$L_{cls} = L_s + L_{emb} \tag{15}$$

Ultimately, the D-D capability of the model is improved by updating the shared encoder and category classifiers with classification losses. The improved ACGAN and AGAN models jointly constitute the complete GAN-TL framework proposed in this study, as shown in Fig. 6.

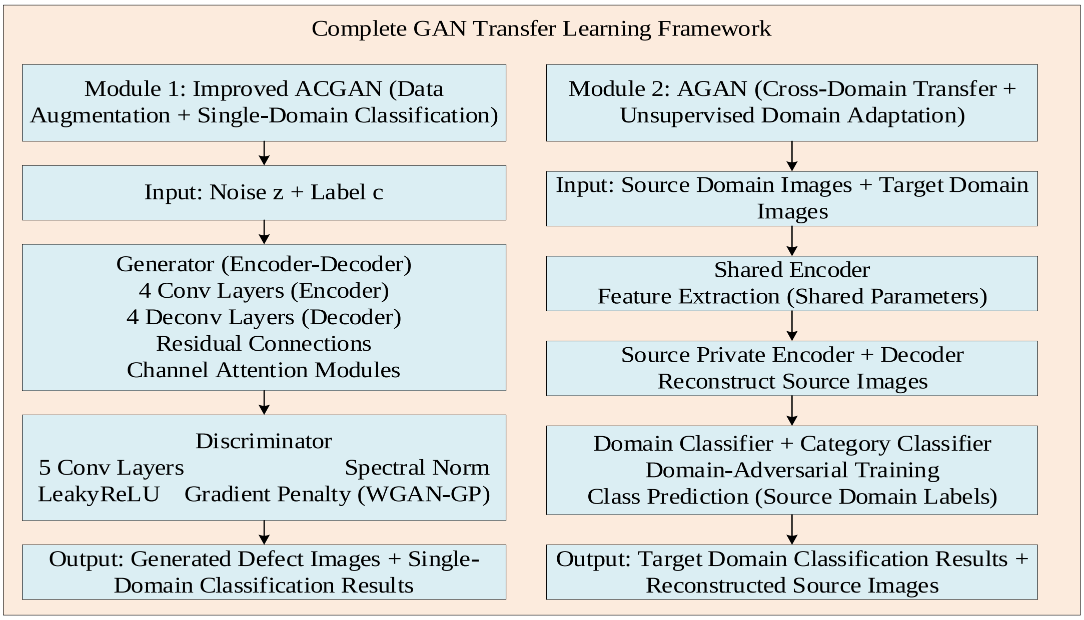


Fig. 6. Network structure diagram of the GAN TL framework

In Fig. 6, the improved ACGAN serves as the data augmentation and domain-specific classification module. Through the adversarial training that incorporates gradient penalty and attention mechanism, it generates high-quality artificial defect samples and completes the defect classification task within a single domain. The AGAN model is built based on unsupervised DA and assumes the function of the cross-domain transfer module. It aligns the feature distributions between the source domain and the target domain through adversarial domain discrimination and feature enhancement, thereby enabling effective defect detection even when there is a scarcity of labeled samples in the target domain. The TL strategy is specifically implemented in the AGAN component: the shared encoder and domain classifier extract domain-invariant features through adversarial training, and the classification loss and reconstruction loss of the source domain data jointly guide the model to generalize to the unlabeled target domain. The improved ACGAN first expands the training set by generating artificial samples to alleviate the problem of data imbalance. Subsequently, when the model is deployed to a new type of metal material or production condition, the AGAN module performs cross-DA.

GAN promotes metal forming defect detection through three approaches: Firstly, the generator synthesizes artificial images that are similar to the distribution of real defect samples, expanding the training set size and alleviating the problems of insufficient defect data annotation and unbalanced categories. Secondly, the discriminator is forced to learn feature representations with strong discriminative power during the adversarial training process. These features can effectively distinguish between normal and defective surfaces and indirectly serve the defect classification task. Thirdly, the anomaly detection mechanism based on reconstruction error - when the reconstruction difference between the input image and the generated image exceeds the set threshold, it is determined as an abnormal area, thereby achieving unsupervised localization and recognition of unseen defect types.

4. Model Performance Testing

4.1. Experimental Setup and Basic Information

The study utilized the dataset of surface defects in aluminum casting profiles and the publicly available dataset of defects in NEU hot-rolled steel strips from Northeastern University to conduct experimental verification of the improved ACGAN algorithm. This algorithm was also compared and tested with Support Vector Machine (SVM), ResNet50, traditional ACGAN, and Wasserstein GAN (WGAN) algorithms. The aluminum casting profile surface defect dataset was collected from industrial images and comprehensively included various typical processing defects such as scratches, ink spots, surface impacts, and surface stains. The dataset had a sufficient total sample quantity. During the experimental stage, 100 images of each type of defect were randomly selected from this dataset as independent test sets. The images were captured using an industrial face array camera combined with a circular lighting device, with a resolution of 2560×1920 pixels. The collection scene was based on the aluminum casting profile automated production line, with the equipment fixed on the side of the production line. The color temperature was kept constant at 5500K to reduce interference from metal surface reflection and uneven illumination. The images were in color, and the defect samples had natural class imbalance characteristics, with significant differences in defect morphology and texture scale. The NEU hot-rolled steel strip dataset was an industrial grayscale image dataset consisting of 1800 images, including six typical surface defects of hot-rolled steel strips: oxide scale inclusions, patches, cracks, pitting surfaces, inclusions, and scratches. Each type of defect was strictly allocated 300 samples. Each image had a standard resolution of 200×200 pixels. All images were collected from the real hot-rolled steel production line. Due to the influence of line illumination fluctuations and slight changes in steel material, the dataset had the characteristics of large differences in defect appearance within classes and high similarity in defect features between classes. All samples were manually accurately labeled, and they could simultaneously support defect classification and feature learning tasks. The experiment employed stratified random sampling to divide each dataset into a training set and a test set, with a ratio of 80% to 20%. This ensured that the proportions of various defects in the training set and test set were consistent with those in the original dataset. In the five-fold cross-validation, the training set was further divided into five mutually exclusive subsets (each fold containing 16% of the total number of samples of each type), and one of these folds was selected as the validation set while the remaining four folds were used as the training subsets.

This process was repeated five times. The final performance was calculated as the average value of the five-fold cross-validation. For 10 repetitions of the experiment, each repetition involved a new 80/20 stratified random division and five-fold cross-validation, and a different random seed (with values ranging from 1 to 10) was used for each repetition to evaluate the model's stability under different data divisions. The final reported accuracy rate was the average ± standard deviation of the 10 repetitions. The dataset of surface defects in aluminum profile processing (surface scratches Sc, surface abrasions Br, surface stains St, ink dots In) contained approximately 550 samples for each category. The experiment adopted a balanced partition strategy, randomly selecting 450 images from each category as the training set, and approximately 100 images remaining as the test set. To comprehensively evaluate the model's performance on various types of defects, a balanced training and testing sample division was adopted (450 samples for each defect category for training, and 100 samples for testing) to avoid the interference of class imbalance on performance evaluation.

All model training and testing experiments were conducted based on a unified hardware platform. The specific hardware configuration was an Intel Core i9-12900K central processor, 128 GB of RAM, and a single NVIDIA RTX 3090 graphics card. The experimental environment was equipped with the Ubuntu 20.04 operating system, the PyTorch deep learning framework, and the CUDA 11.4 parallel computing library. During the training phase, the total training time required for the improved ACGAN model to complete all iterations and converge was approximately 14.6 hours. Under the same experimental conditions, the training times for the traditional ACGAN and ResNet50 models were 12.1 hours and 9.3 hours. This model, due to the introduction of residual structures, attention mechanisms, and gradient penalty modules, had an increase in network parameters and computational load, resulting in a slight increase in training time. However, it remained within the reasonable range for industrial model iterative optimization. In the inference stage, for NEU steel defect images with a resolution of 200×200 and aluminum profile surface defect images with a resolution of 2560×1920, the average inference time for a single image of the improved ACGAN model was 18 ms and 47 ms, respectively. The inference speed could stably meet the sampling and online detection cycle requirements of industrial production lines. In terms of computational load, the total number of parameters in the improved ACGAN model was approximately 24.6M (million), and the number of floating-point operations for forward inference of a single image (with a resolution of 256×256) was approximately 8.7G FLOPs. The parameter count of the baseline ACGAN model was 18.2M, and the FLOPs were approximately 6.4G. The improved ACGAN achieved a good balance between computational efficiency and detection accuracy and was suitable for industrial defect detection scenarios that required certain real-time performance.

To evaluate the convergence and discrimination ability of the model training process, the loss curves of the generator and discriminator as well as the ROC curve of the classifier during the improved ACGAN training were recorded, as shown in Fig. 7. Within 200 training iterations, the generator loss decreased from the initial 2.34 to 0.42, while the discriminator loss dropped from 1.87 to 0.38 and then stabilized. The two reached equilibrium approximately after 120 iterations, indicating that the model training process was stable and there was no significant pattern collapse. The area under the ROC curve (AUC) of the improved ACGAN was 0.976, while that of the baseline ACGAN was 0.941. This further verified the superiority of the improved model in discriminating between real samples and generated samples.

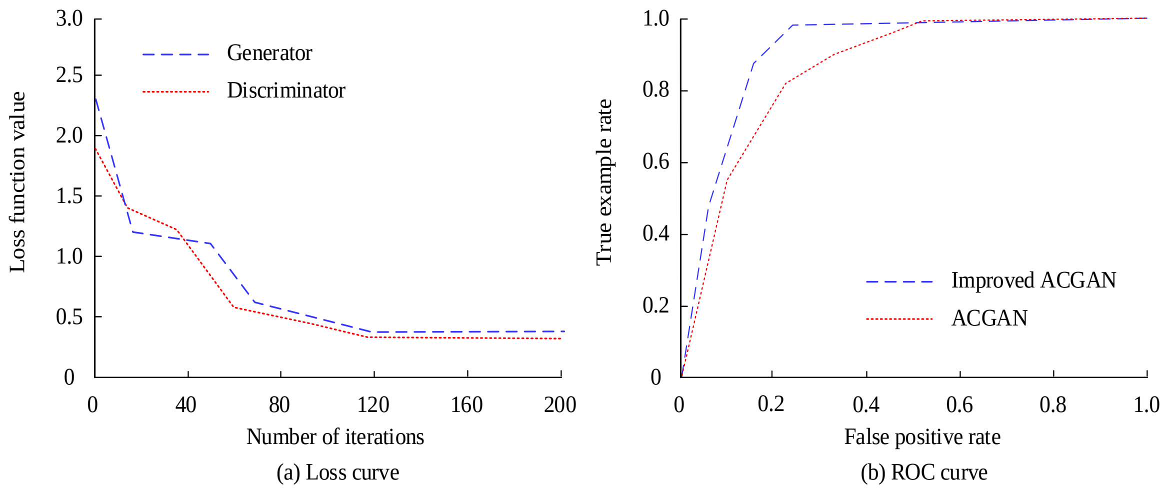


Fig. 7. GAN loss curve and ROC curve

To assess the sensitivity of the gradient penalty coefficient, while keeping other hyperparameters fixed, the gradient penalty coefficient was set to 1, 5, 10, 15, and 20 for the experiments. The results showed that when the gradient penalty coefficient was 10, the model achieved the best classification accuracy (97.85%) and FID value (0.289). When the gradient penalty coefficient was 1, there were slight signs of mode collapse during training, with the accuracy dropping to 92.31%

and the FID rising to 0.412. When the gradient penalty coefficient was 20, the excessive gradient penalty led to slow convergence of the discriminator, with an accuracy of 94.56% and an FID of 0.356. The model performance was relatively stable within the range of 5 to 15 for the gradient penalty coefficient, indicating that the improvement of ACGAN had certain robustness to the value of the gradient penalty coefficient. The gradient penalty coefficient of 10 was the best choice based on the experimental results.

In terms of real-time industrial deployment, the single-image inference time of the improved ACGAN has been reduced to 18ms (NEU 200×200 pixels) and 47ms (aluminum profile 2560×1920 pixels), with a parameter size of 24.6M and FLOPs approximately 8.7G (256×256 input). Compared to the baseline ACGAN (18.2M, 6.4G), the parameter size and computational load have slightly increased, but the inference speed still met the production line cycle requirements (typically ≤ 100ms/image). If deployed on edge devices with limited computing power, model pruning or knowledge distillation could be used to compress the inference time to within 25ms, with the accuracy loss controlled within 1.5%.

4.2. Ablation Study

To clarify the independent contributions of each improvement to the model performance, an ablation experiment was designed. Taking the standard ACGAN as the baseline model, three improvements were introduced successively: the residual block, the channel Attention mechanism, and the Gradient Penalty. On the aluminum profile surface defect dataset (200 training samples for each class), a five-fold cross-validation was employed to record the impact of each improvement on the classification accuracy. The ablation experiment results are shown in Table 1.

Table 1. Influence of different improvement modules on the classification accuracy of the model

Model configuration	Average accuracy rate (%)	Improvement over baseline (%)
Baseline ACGAN	92.43	/
Baseline + Residual Module	94.56	+2.13
Baseline + Residual Module + Attention Mechanism	96.31	+3.88
Baseline + Residual Module + Attention Mechanism + Gradient Penalty (Complete Improvement of ACGAN)	97.85	+5.42

From Table 1, the residual module contributed 2.13% to the accuracy improvement. The attention mechanism further increased the accuracy by 1.75% on top of this, while the gradient penalty contributed 1.54% to the improvement. The combined effect of these three improvements raised the model's accuracy from 92.43% to 97.85%, verifying that each modification had a positive impact on the final performance.

4.3. Comparative Analysis

The results of the relevant experimental tests are provided in Table 2. The improved ACGAN and ACGAN were tested at various training sizes, with the image set used for experimental testing being fixed. Both models were put through ten experiments, with the average value being used as the final outcome.

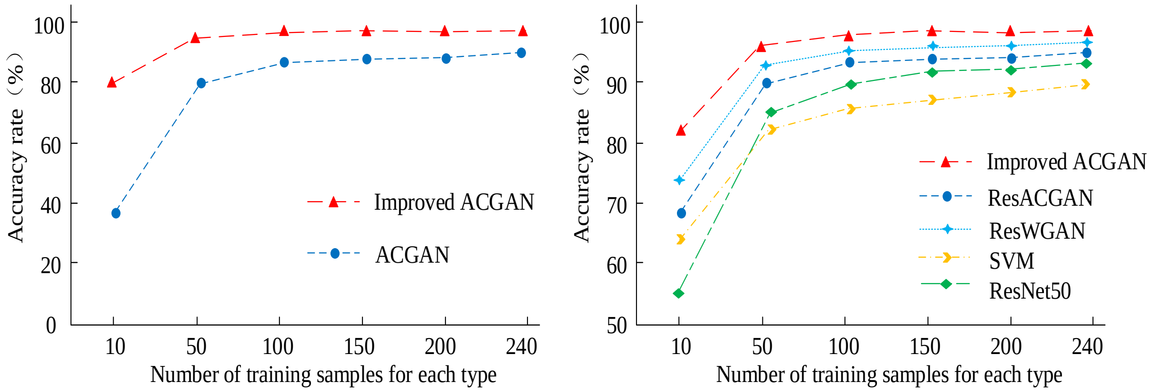
In Table 2, the accuracy of Improved ACGAN was higher than that of ACGAN under different training sizes, and the average accuracy of Improved ACGAN was 95.98%. Even when there were only 10 training samples per category, the average accuracy of improved ACGAN was as high as 83.74%. As the number of training samples continued to decrease, both methods, Improved ACGAN and ACGAN, showed a decreasing trend, but the recognition accuracy of Improved ACGAN decreased significantly more slowly, and the accuracy of Improved ACGAN increased by 41.74% compared to ACGAN. This result also verified the adaptability and superiority of Improved ACGAN on small sample data. Under the condition of 200 training samples for each category, the standard deviation of the accuracy rate of the 10 experiments of the improved ACGAN was 1.24%, while that of the baseline ACGAN was 2.87% and that of ResNet50 was 3.42%. This indicated that the improved

ACGAN had more stable performance in multiple repeated experiments. Further, a paired t-test was used to conduct a significance test on the 10 accuracy rate differences between the improved ACGAN and the baseline ACGAN, and the calculated p-value was 0.003 (<0.05), indicating that the performance difference between the two was statistically significant. The paired t-test p-value between the improved ACGAN and ResNet50 was 0.001, also showing a significant difference.

Table 2. Improved accuracy of ACGAN and ACGAN

Sample size	ACGAN average test accuracy %	Improved ACGAN average test accuracy %	Growth rate %
240	94.00	99.53	5.53
200	92.43	99.42	6.99
150	91.08	98.68	7.6
100	88.65	98.31	9.66
50	74.37	96.25	21.88
10	40.12	83.74	41.74

To further verify the detection accuracy of the improved ACGAN, it was compared with SVM and ResNet classification methods. The experimental results of each model under various sample sizes and the defect classification results under various methods are shown in Fig. 8, where the average of ten times is taken as the final result in order to avoid the influence of the initial parameters on the DL model.



（a）Experimental results of each model under different sample sizes

（b）Results of defect classification under different methods

Fig. 8. Experimental results of each model under different sample size and defect classification results under different methods

In Fig. 8, the performance of the GAN-based optimization strategy consistently outperformed SVM and ResNet50 at various training volumes, demonstrating how sample generation by generative pair networks could increase the model's classification accuracy in a small-sample setting. The model proposed by the study still achieved the best results, verifying the improvement of the GAN network based on the regret minimizing gradient penalty term. The experimental comparison showed that the improved ACGAN model proposed by the study was able to obtain better D-D capability with a limited number of samples. Due to the imbalance of categories in the data of metal surface defects, relying solely on accuracy to evaluate the model performance might not be comprehensive enough. Therefore, in addition to accuracy, macro-average precision, recall rate, F1 score, and AUC-ROC value were also adopted as evaluation indicators. Under the condition of 50 training samples for each category, the macro-average precision of the improved ACGAN was 94.82%, the macro-average recall rate was 93.57%, the macro-average F1 score was 94.19%, and the AUC-ROC value was 0.976. The corresponding indicators of the baseline ACGAN were 88.31%, 86.74%, 87.52%, and 0.941. The above results indicated that the comprehensive performance of the improved ACGAN in all categories was superior to that of the baseline model, verifying its effectiveness and robustness in the scenario of category imbalance.

To test the performance of adversarial generation and unsupervised DA based defect classification models, the study set up two sets of domain migration experiments with defects on different metal surfaces: cast iron - aluminum profile, and aluminum profile - cast iron. Unsupervised DA based Convolutional Neural Network (CNN), Direct Digital Control (DDC) and Deep Adaptation Networks (DAN) models were compared under the same base model and experimental conditions.

Table 3. Experimental comparison of domain migration of metal surface defects

Method	Cast iron - aluminum profile	Aluminum profile - Cast iron	Average accuracy
AGAN	77.14	86.11	81.63
CNN	60.50	76.00	68.25
DDC	70.00	81.68	75.84
DAN	75.56	86.11	80.84

In Table 3, the average accuracy was only 68.25% when trained using CNN alone, while the recognition accuracy was improved using the unsupervised model based on conventional domain alignment optimization, which could reach 75.84% and 80.84%, and the AGAN model described in the study, which had an average prediction accuracy of 81.63%, which improved the unsupervised DA algorithms when compared with other accuracy of D-D on metal surfaces.

To further verify the superiority of the improved ACGAN model in the task of metal surface defect detection, it was compared with four current mainstream deep learning models: Vision Transformer (ViT) as the benchmark model for image classification in the Transformer architecture [20], YOLOv8 as the object detection model to compare the joint performance of defect location and classification [21], EfficientNet-B3 as an efficient lightweight CNN classification network [22], and Diffusion Model for comparing the quality of defect sample generation and anomaly detection performance [23]. All the comparison models used the same data division (200 training samples for each category) and the five-fold cross-validation strategy and were trained and tested on the same hardware platform. ViT adopted the pre-trained ViT-B/16 architecture, YOLOv8 used the YOLOv8n version, EfficientNet-B3 used ImageNet pre-trained weights for fine-tuning, and the diffusion model used the Denoising Diffusion Probabilistic Model (DDPM) framework and trained for 200 rounds on the defect generation task. To comprehensively evaluate the performance of the improved ACGAN model across different task dimensions, ViT-B/16 (image classification), EfficientNet-B3 (image classification), YOLOv8n (object detection), and the diffusion model based on DDPM (anomaly detection and generation) were selected as comparison methods. Table 4 summarizes the performance of each method on the aluminum profile surface defect dataset.

Models	Improved ACGAN	ACGAN	ViT-B/16	EfficientNet-B3	YOLOv8n	DDPM
Task type	Classification + Generation	Classification + Generation Classification	Classification	Classification	Detection (location + classification)	Generation + anomaly detection
Accuracy (%)	97.85	92.43	93.76	94.51	/	/
Precision (%)	96.71	91.52	92.87	93.64	/	/
Recall rate (%)	96.14	90.84	92.22	92.95	/	/
F1(%)	96.42	91.18	92.54	93.29	/	/
ROC-AUC	0.976	0.941	0.958	0.963	/	/
Inference time (ms/ Graph)	38	42	52	35	28	180 (generation) / 35 (testing)
FLOPs(G)	8.7	6.4	17.6	1.9	8.7	/
Parameter quantity (M)	24.6	18.2	86.6	12.3	3.2	35.7

In terms of classification tasks, the accuracy (97.85%), precision (96.71%), recall (96.14%), F1 value (96.42%), and ROC-AUC value (0.976) of the improved ACGAN were all superior to those of ViT-B/16 and EfficientNet-B3. The confusion matrix showed that the classification accuracy of the improved ACGAN in each defect category was significantly higher than that of the baseline ACGAN, especially in the surface stains and embedded objects categories, where the improvement was most notable. In terms of computational efficiency, the inference time of the improved ACGAN was 38ms per image, with FLOPs of 8.7G and parameters of 24.6M. The inference speed was better than ViT-B/16 (52ms, 17.6G, 86.6M), although it was slower than EfficientNet-B3 (35ms, 1.9G, 12.3M) and YOLOv8n (28ms, 8.7G, 3.2M). However, it achieved a better balance between accuracy and efficiency, making it suitable for industrial scenarios with high requirements for detection accuracy. In terms of localization performance, the mAP@0.5 of YOLOv8n was 89.34%, while the average IoU obtained by thresholding the heat map of the improved ACGAN was 67.3%. In terms of generation quality, the FID value of the diffusion model (0.267) was slightly better than that of the improved ACGAN (0.289), but the improved ACGAN integrated generation, classification, and localization functions in the same framework, which had better system integration and inference efficiency in practical industrial deployment. The improved ACGAN performed exceptionally well in classification tasks.

To evaluate the performance of the model under extremely sparse annotation conditions, training scenarios were set on the aluminum profile dataset where each category was provided with only 1, 5, 10, and 20 labeled samples, and the remaining training images were regarded as unlabeled. The improved ACGAN, leveraging pre-training and data augmentation for generation, achieved an average accuracy of 62.4% in 1-shot conditions, 78.6% in 5-shot conditions, 83.7% in 10-shot conditions, and 89.2% in 20-shot conditions. While the baseline ACGAN achieved only 38.7%, 55.3%, 64.1%, and 74.5% respectively under the same conditions. The results indicated that the improved model could maintain high performance even when the number of labeled samples was very small. Its advantages mainly stemmed from the general feature priors provided by the pre-training initialization and the expanded training data of the generator.

4.4. Application Results

An example of the comparison between the steel surface defect samples generated by the improved ACGAN model and the real samples is shown in Fig. 9.

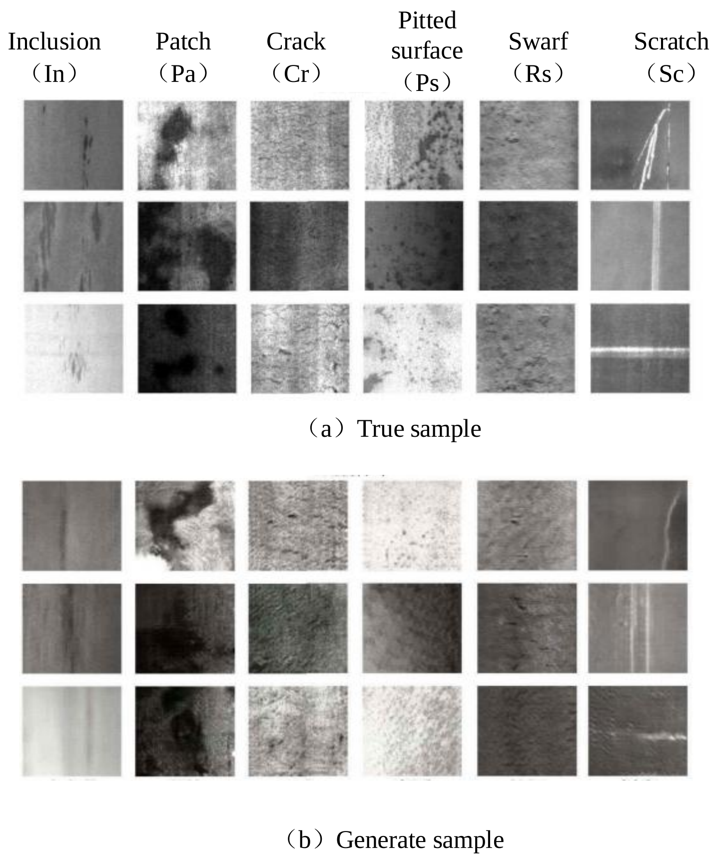


Fig. 9. Improved ACGAN to generate a comparison diagram of the results of steel surface defect samples

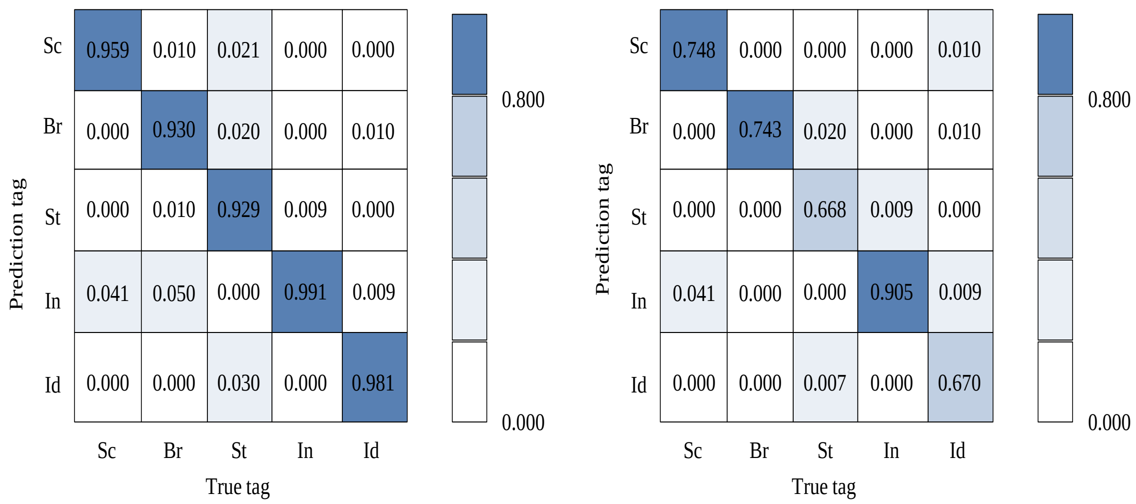
Fig. 9 presents a visual display of both real and generated defect samples. An analysis was conducted on 6 types of steel surface faults: The real samples clearly demonstrated the original industrial morphological characteristics of inclusions, patches, cracks, etc. The model-generated samples highly replicated the textures and key features of various faults, while preserving the internal morphological differences, which could effectively expand the fault data set and optimize the stability of the industrial quality inspection process in identifying various surface faults. To quantitatively evaluate the difference in generation quality between Improved ACGAN and other methods, Mode Score (MS) and Frechet Inception Distance (FID) metrics were used to evaluate the quality of generated samples. Table 5 compares the average MS and FID values of the samples produced by various models with ACGAN, ResACGAN, and ResWGAN, respectively.

Table 5. Comparison of average MS and FID of samples generated by different models

Data set	Method	MS	FID
NEU Data collection of steel surface defects	Improved ACGAN	1.386	0.289
	ResWGAN	1.374	0.311
	ResACGAN	1.361	0.317
	ACGAN	1.336	0.342

In Table 5, the MS scores of the improved ACGAN were 1.386, which were all higher than those of the other models, and the improved method also achieved a certain level of improvement from the FID point of view, which proved that the improved ACGAN model achieved a certain degree of fitting to the real data. Furthermore, the Learned Perceptual Image Patch Similarity (LPIPS) and the variance ratio of the feature space were introduced to quantify the intra-class diversity of the generated samples. On the NEU steel surface defect dataset, the average LPIPS value of the improved ACGAN-generated samples was 0.347, while that of the baseline ACGAN was 0.291 and that of ResWGAN was 0.312. This indicated that the improved ACGAN generated samples of the same defect type with more diverse textures and morphologies. In terms of the variance ratio of the feature space, the ratio of the improved ACGAN was 0.94, that of the baseline ACGAN was 0.78, and that of ResWGAN was 0.85. This showed that the feature distribution of the improved ACGAN-generated samples was more similar to that of the real samples, and better retained the intrinsic diversity of various defects.

To better show the D-D effect of Improved ACGAN on real industrial data, the study introduced a multi-classification confusion matrix to further quantify the detection effect, and the average confusion matrix result plots of Improved ACGAN and ACGAN on the aluminum machined surface defects dataset are shown in Fig. 10.



（a）Improved ACGAN （b）ACGAN

Fig. 10. The average confusion matrix of ACGAN and ACGAN defect data sets on aluminum machined surfaces is improved

In Fig. 10, the horizontal axis of the confusion matrix represented the true defect category, and the vertical axis represented the predicted defect category by the model. The values on the diagonal represented the classification accuracy of the corresponding categories. The improved ACGAN achieved diagonal accuracy rates of 0.959, 0.930, 0.929, 0.991, and 0.981 for the five types of defects (Sc, Br, St, In, Id), which were significantly higher than the original ACGAN's rates of 0.748, 0.743, 0.663, 0.905, and 0.670. Particularly, the recognition accuracy for St and Id defects increased significantly. Both models could maintain a high recognition level for the in defect. The original ACGAN was prone to misclassification between adjacent categories in the other categories. The improvement scheme effectively reduced the probability of misclassification across categories.

The improved model could be deployed in industrial production lines for real-time surface defect detection of metal workpieces. Production lines suffered from dynamic lighting variations, and reflections of aluminum profiles and steel change surface brightness, leading to obscured defect textures and blurred defect boundaries. Industrial cameras had offset shooting angles and focal lengths, bringing motion blur, while circuit transmission and workshop electromagnetic interference introduced Gaussian noise and salt-and-pepper noise. This model adopted a multi-scale adaptive brightness equalization module to correct image grayscale distribution, an angle-invariant feature extraction branch to reduce feature distortion, and noise reduction convolution layers to eliminate motion blur and electromagnetic noise, so it could extract core defect features steadily against imaging interferences. Its lightweight backbone cut floating-point computations. The inference time for single images fit production line tact time, matching continuous workpiece conveyance and avoiding workpiece accumulation and detection lag. The model delivers reliable performance for industrial real-world deployment.

To evaluate the robustness of the model under different manufacturing conditions, three sets of generalization tests were designed on the aluminum profile surface defect dataset: Firstly, the training set (obtained under uniform lighting and fixed angles) was compared with two independently collected test sets, which respectively simulated the working conditions of light intensity fluctuation (±30%) and camera angle deviation (±15°). Secondly, Gaussian noise of different intensities (σ = 5, 10, 20) and motion blur (kernel size 3×3, 5×5, 7×7) were added to the test images to simulate the common image quality degradation in industrial environments. Thirdly, defect samples from aluminum profiles of different batches and different production lines were selected as the second independent test set to evaluate the model's adaptability to material surface differences between batches. The experimental results showed that the accuracy of the improved ACGAN decreased from the baseline 97.85% to 94.23% under light intensity fluctuation conditions, to 92.67% under angle deviation, to 89.41% after adding Gaussian noise (σ = 20), and to 90.76% after motion blur (7×7). The accuracy on the independent test set across batches was 91.54%. As a comparison, the baseline ACGAN's accuracy under the same conditions decreased to 86.34%, 84.72%, 78.56%, 80.13%, and 83.27% respectively. These results verified that the improved model had stronger anti-interference ability and generalization performance in different manufacturing environments.

To quantify the model's sensitivity to changes in lighting and imaging conditions, an interference experiment was designed on the aluminum profile test set. When the illuminance fluctuated within the range of 1000lx ± 30%, the accuracy rate remained above 96%. When it deviated by ±50%, it dropped to 91.7%. When the angle between the light source and the optical axis was within 30°, the accuracy rate remained above 96.8%, dropping to 93.4% at 45° and 88.2% at 60°. The exposure time was stable within the range of 200μs to 500μs, and when the gain exceeded 12dB, the accuracy rate decreased from 95.6% to 90.3%. The baseline ACGAN had a 6 to 11 percentage point higher accuracy rate reduction under various conditions compared to the improved model. The model had good robustness to moderate lighting fluctuations, but was more sensitive to extreme underexposure, overexposure, and large-angle tilted lighting. Industrial deployment required either stable lighting or the introduction of an adaptive brightness correction module.

To evaluate the model's ability to detect unseen defect categories, a zero-shot generalization experiment was designed. One category was excluded from the training process in each of the four defect types, and the remaining three categories were used to train the model. The model's ability to recognize the excluded category was tested. The average recall rate of the four experiments was 37.6%, and the F1 value was 35.2%. After performing rapid fine-tuning with a small number of labeled samples for new categories (1-shot, 5-shot, 10-shot), the average recall rates increased to 52.3%, 71.8%, and 83.6%, respectively. The results indicated that the model's ability to recognize completely new defects under zero-shot conditions was limited, but it could quickly adapt with the guidance of a small number of new category samples and had good scalability.

5. Conclusion

DFDMM guaranteed product quality and safety. Traditional methods relied on massive labeled samples and manual feature design, limiting practical use. This work combined GAN and TL for DFDMM. The pre-trained GAN transferred knowledge to target domains for performance improvement. Under the extremely small sample condition where each category contained only 10 training samples, the average accuracy of the improved ACGAN reached 83.74%. When the training sample size for each category was 240, the average accuracy could reach 99.53%, and the overall average accuracy of the improved ACGAN in the experiment was 95.98%. It performed well on imbalanced aluminum profile defect data, with over 98% accuracy on scratches and ink dots In. AGAN obtained 81.63% accuracy. The multi-task sharing mechanism and DA boosted overall performance and cross-dataset adaptability. This model could be embedded into industrial visual inspection systems for efficient defect detection. However, the research also had limitations. The model performance was highly dependent on the distribution and characteristics of the adopted datasets, and its detection accuracy might degrade obviously when applied to practical production scenarios with more complex working conditions. For instance, in a dynamic industrial environment, the model might be disturbed, affecting its performance. Future research can further optimize the model structure and algorithms to adapt to more complex defect detection tasks, such as extending to 3D defect detection. Furthermore, by integrating other deep learning techniques, such as reinforcement learning or graph neural networks, the performance of the model may be further enhanced.

Funding

The research is supported by National Natural Science Foundation of China, Research on the Anti-icing Mechanism and Anti-icing Characteristics of Composite Anti-icing Coatings Based on Surface Engineering and Microelectrode Transient Heat Release Technology, No.: 52205238.

References

[1] Pavan Kumar MR, Jayagopal P. Generative adversarial networks: a survey on applications and challenges. Int J Multimed Inf Retr. 2021;10(1):1-24. https://doi.org/10.1007/s13735-020-00196-w
[2] Tan J, Liao X, Liu J, Cao Y, Jiang H. Channel attention image steganography with generative adversarial networks. IEEE Trans Netw Sci Eng. 2022;9(2):888-903. https://doi.org/10.1109/TNSE.2021.3139671
[3] Jain S, Seth G, Paruthi A, Soni U, Kumar G. Synthetic data augmentation for surface defect detection and classification using deep learning. J Intell Manuf. 2022;33(4):1007-1020. https://doi.org/10.1007/s10845-020-01710-x
[4] Luo Q, Fang X, Liu L, Yang C, Sun Y. Automated visual defect detection for flat steel surface: a survey. IEEE Trans Instrum Meas. 2020;69(3):626-644. https://doi.org/10.1109/TIM.2019.2963555
[5] Ren Z, Fang F, Yan N, Wu Y. State of the art in defect detection based on machine vision. Int J Precis Eng Manuf-Green Technol. 2022;9(2):661-691. https://doi.org/10.1007/s40684-021-00343-6
[6] Zhou X, Hu Y, Wu J, Liang W, Ma J, Jin Q. Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT. IEEE Trans Ind Inform. 2023;19(1):570-580. https://doi.org/10.1109/TII.2022.3170149
[7] Lei X, Sun L, Xia Y. Lost data reconstruction for structural health monitoring using deep convolutional generative adversarial networks. Struct Health Monit. 2021;20(4):2069-2087. https://doi.org/10.1177/1475921720959226
[8] Cai Z, Xiong Z, Xu H, Wang P, Li W, Pan Y. Generative adversarial networks: a survey toward private and secure applications. ACM Comput Surv. 2021;54(6):1-38. https://doi.org/10.1145/3459992
[9] Hu C, Wang Y. An efficient convolutional neural network model based on object-level attention mechanism for casting defect detection on radiography images. IEEE Trans Ind Electron. 2020;67(12):10922-10930. https://doi.org/10.1109/TIE.2019.2962437

[10] Rajabi A, Garibay O. TabFairGAN: fair tabular data generation with generative adversarial networks. Mach Learn Knowl Extr. 2022;4(2):488-501. https://doi.org/10.3390/make4020022
[11] Krause J, Grabsch HI, Kloor M, Jendrusch M, Echle A, Buelow RD, et al. Deep learning detects genetic alterations in cancer histology generated by adversarial networks. J Pathol. 2021;254(1):70-79. https://doi.org/10.1002/path.5638
[12] Baumgartl H, Tomas J, Buettner R, Merkel M. A deep learning-based model for defect detection in laser-powder bed fusion using in-situ thermographic monitoring. Prog Addit Manuf. 2020;5(3):277-285. https://doi.org/10.1007/s40964-019-00108-3
[13] Yang Y, Liu J, Huang S, Wan W, Wen W, Guan J. Infrared and visible image fusion via texture conditional generative adversarial network. IEEE Trans Circuits Syst Video Technol. 2021;31(12):4771-4783. https://doi.org/10.1109/TCSVT.2021.3054584
[14] Jabbar A, Li X, Omar B. A survey on generative adversarial networks: variants, applications, and training. ACM Comput Surv. 2021;54(8):1-49. https://doi.org/10.1145/3463475
[15] Jing J, Wang Z, Rätsch M, Zhang H. Mobile-Unet: an efficient convolutional neural network for fabric defect detection. Text Res J. 2022;92(1-2):30-42. https://doi.org/10.1177/0040517520928604
[16] Tabernik D, Šela S, Skvarč J, Skočaj D. Segmentation-based deep-learning approach for surface-defect detection. J Intell Manuf. 2020;31(3):759-776. https://doi.org/10.1007/s10845-019-01476-x
[17] Su B, Chen H, Chen P, Bian G, Liu K, Liu W. Deep learning-based solar-cell manufacturing defect detection with complementary attention network. IEEE Trans Ind Inform. 2021;17(6):4084-4095. https://doi.org/10.1109/TII.2020.3008021
[18] Hu G, Huang J, Wang Q, Li J, Xu Z, Huang X. Unsupervised fabric defect detection based on a deep convolutional generative adversarial network. Text Res J. 2020;90(3-4):247-270. https://doi.org/10.1177/0040517519862880
[19] Aldausari N, Sowmya A, Marcus N, Mohammadi G. Video generative adversarial networks: a review. ACM Comput Surv. 2022;55(2):30. https://doi.org/10.1145/3487891
[20] Ayon STK, Siraj FM, Uddin J. Steel surface defect detection using learnable memory vision transformer. Comput Mater Contin. 2025;82(1):499-520. https://doi.org/10.32604/cmc.2025.058361
[21] Wang Y, Zhang K, Wang L, Wu L. An improved YOLOv8 algorithm for rail surface defect detection. IEEE Access. 2024;12:44984-44997. https://doi.org/10.1109/ACCESS.2024.3380009
[22] Zhang H, Yang Z, Lei N. Defect recognition of solar panel in EfficientNet-B3 network based on CBAM attention mechanism. In: 2024 International Conference on Generative Artificial Intelligence and Information Security (GAIIS 2024); 2024. p. 349-352. https://doi.org/10.1145/3665348.3665407
[23] Zhou X, Zhang Y, Ren Z, Mi T, Feng K, Zhou S, et al. DiffDD: a surface defect detection framework with diffusion probabilistic model. Adv Eng Inform. 2024;62:102637. https://doi.org/10.1016/j.aei.2024.102637