Unsupervised Anomaly Detection of Industrial Images Based on Dual Generator Reconstruction Networks

Cong Gu

School of Mathematics and Information Science, Zhongyuan University of Technology,
Zhengzhou Henan 450007, China
Zhengzhou Key Laboratory of Image Processing and Optoelectronic Information Technology,
Zhengzhou Henan 451191, China

Siyv Ren, Qiqiang Duan

School of Mathematics and Information Science, Zhongyuan University of Technology,
Zhengzhou Henan 450007, China

Corresponding author: gucong@zut.edu.cn

At present, deep learning techniques are increasingly utilized in computer vision and anomaly detection. To address the limitations of inadequate reconstruction capability and subpar performance in reconstruction-based anomaly detection, this study enhances the existing algorithm and introduces an unsupervised anomaly detection of industrial images algorithm based on dual generator reconstruction networks-DGRNet. The network consists of two generators and a discriminator, introducing a widely recognized denoising diffusion probabilistic model (DDPM) as one of the generators, an autoencoder (AE) as the other generator, and a decoder as the discriminator. The model is tested on the MVTec AD dataset, and in the case of no additional training data, the anomaly detection AUC result of DGRNet exceeds the baseline method based on reconstruction by 19.6 percentage points. The experimental results show that DGRNet can improve the detection performance in the anomaly detection algorithm based on unsupervised and reconstructed networks.

KEYWORDS: Anomaly detection, unsupervised learning, Denoising Diffusion Probabilistic Model (DDPM), reconstruction-based network, autoencoder.

1. Introduction

Anomaly detection is an important branch of machine learning, a technology for detecting abnormal situations and mining illogical data and is widely applied in different fields such as credit card fraud, insurance or healthcare, cybersecurity intrusions, safety-critical systems, industrial big data, abnormal behaviors, image, and video processing. In image anomaly detection, there are mainly three aspects of applications, namely defect detection, medical image analysis, and hyperspectral image processing. Early anomaly detection was mostly applied in the field of data mining, and relatively traditional algorithms were used. Traditional algorithms are generally supervised learning methods, which can be divided into three categories: similarity measure-based, statistical, and probability-based, and linear model-based [9].

In recent years, with the rapid development of deep learning methods, anomaly detection has been introduced to the field of image processing to solve problems such as target detection which has rare sample [11]. In anomaly detection tasks, it is usually difficult to obtain well-labeled anomalous samples, so the research on unsupervised anomaly detection in deep learning had received more attention. Among unsupervised anomaly detection algorithms, reconstruction-based algorithm is a natural idea. This kind of algorithms generally include algorithms based on Autoencoders (AE) [16] and Generative Adversarial Networks (GAN) [7], as well as algorithms that combine AE and GAN. In the case of AE-based algorithms, during the training stage, only normal samples are used and during the inference stage, the difference between the normal image and its reconstructed image is very small, while the difference between the abnormal image and its reconstructed image is large, and the size of the difference can be used to judge whether the image is abnormal. Mei et al. [12] reconstructed images by dividing them into blocks and used denoising AE to locate anomalies in texture images. GAN, as a generation network, can reconstruct clearer images in algorithms based on GAN. AnoGAN was proposed by Schlegl et al. [19], which directly utilizes GAN iterative optimization to reconstruct images. However, the main problem of this method is its time-consuming during model inference. To solve this issue, EBGAN [23] and other models were proposed. Xiao Du et al. introduced a correction branch to modify the original reconstruction results obtained from the reconstructed network during the test to solve the

problem that it is difficult to distinguish small anomalies in the reconstruction-based network [5]. Zhou et al. proposed to reconstruct the image by leveraging the structure-texture correspondence [24]. Farady et al. proposed Hierarchical Image Transformation and Multi-level Features (HIT-MiLF) modules for an anomaly detection network to adapt to perturbations from novelties in testing images [6]. In the algorithms combining AE and GAN, a discriminator is added after the AE, which is used to distinguish between the reconstructed image and the input image. During training, adversarial training is adopted to enhance the reconstruction ability. The ALOOC model proposed by Sabokrou et al. [17], the ALOOC model consists of two parts: a denoising auto-encoder and a convolutional neural network classifier. These two networks work together to improve the performance of anomaly detection. The GANomaly [1] network structure includes a generator and a discriminator, where the generator is in the form of an encoder-decoder-encoder. It judges whether the input image is an abnormal sample by comparing the potential information of the input image and the reconstructed image. Building on GANomaly, the Skip-GANomaly [2] model was proposed, which adds a skip connection structure to enhance the model's reconstruction ability. However, this model still has the potential to miss detections. When performing surface defect detection, some small-scale defects are easily reconstructed, resulting in small abnormal scores, thereby causing missed detections.

For the past few years, diffusion model (Denoising Diffusion Probabilistic Model, DDPM) is emerging as an unsupervised generative model is a type of unsupervised generative mode. It was first proposed in 2015, but only proposed adding noise to data through a Markov chain during the diffusion process, as well as learning how to reconstruct from noise to data samples [20]. In 2020, DDPM was implemented and gradually became a new hotspot in the generation field [8], demonstrating its powerful generation capabilities, and Dhariwal proved that DDPM can surpass GAN [4]. In the recent wave of AI painting in the field of computer vision, DDPM played a crucial role. In popular AI painting models, OpenAI's DALL·E2 [13] can generate high-definition images just by inputting text; Google's Image [18] is also an AI painting model that can generate images that meet requirements through text descriptions; Stability AI's Stable Diffusion [14] is also

such a model, and the generated images can already be used commercially. These models all demonstrate the powerful image generation capabilities of DDPM. In the field of anomaly detection, AnoDDPM [22] improved DDPM by replacing Gaussian noise with multi-scale simplex noise to capture abnormal areas, without the need for stable training on large datasets. Although the algorithm based on DDPM is not perfect for anomaly detection tasks, DDPM's powerful generative ability is helpful in the field of anomaly detection. We can utilize the powerful generative ability of DDPM to improve the reconstruction ability of reconstruction-based networks in anomaly detection.

The proposed algorithm in this paper is an unsupervised anomaly detection algorithm based on reconstruction. To enhance the reconstruction capability of the network model, this paper proposes an unsupervised image anomaly detection algorithm based on a dual-generator reconstruction network DGRNet, and experiments on industrial images. The algorithm includes two generators, a DDPM generator and an AE generator, and improvements are made to both generators. Each generator reconstructs the input image, and the final reconstructed image is obtained by fusing the features of the two generators. Finally, it is input into a discriminator network for anomaly detection.

2. Methods

2.1. Overview

The proposed in this paper aims to enhance the performance of anomaly detection algorithms based on reconstruction. Compared with existing reconstruction-based networks, the network proposed in this paper has two generators. The overall structure of the network includes two generators and a discriminator, one generator is DDPM, another generator is AE, and a decoder as the discriminator. The generation effect of AE on industrial dataset is not very good, especially when conducting surface defect detection. Some small-scale defects are easily reconstructed, resulting in a small anomaly score, thereby causing missed detection. This paper improves AE by adding attention modules in the encoder of AE, and adding a Skip Connection structure between the last layer of the encoder and the decoder. This article takes the popular DDPM in the generation model as another generation branch, the addition of DDPM improves the ability of network generation and the parameterization of the noise addition steps in the DDPM diffusion process is performed. The overall network structure is shown in Figure 1. Images x_0 is put into two genera-

Figure 1

Overall structure diagram of DGRNet
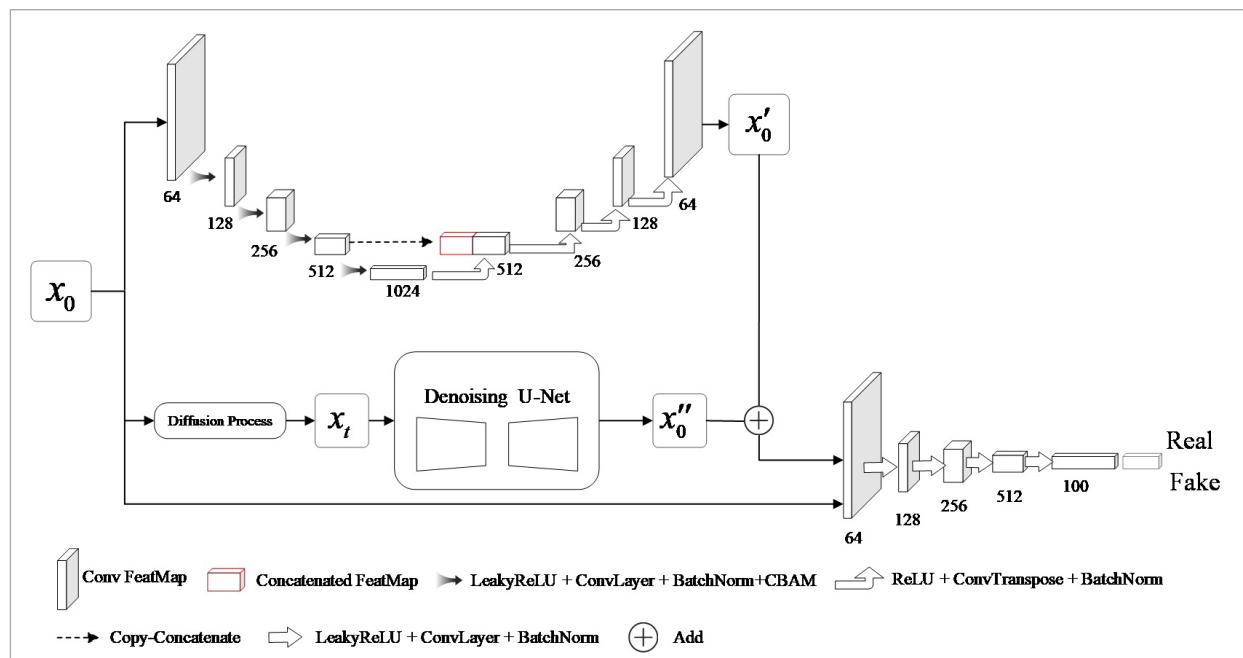
tors, then images x_0' and image x_0'' are reconstructed respectively. Then, the feature fusion of the two reconstructed images yields the final reconstructed image. Finally, the final reconstructed image is input into the discriminator, The latent representation in the discriminator is used for training, and the difference of latent features and the difference between the final reconstructed image and the original image x_0 are used to calculate the anomaly score.

2.2. DDPM Generator

DDPM is a powerful generative model that consists of two processes: forward diffusion process and inverse reconstruction process. The diffusion process refers to the process of adding noise to images, while the reconstruction is the process of denoising images from noisy images to clear original images. In the diffusion process, given the initial data distribution $x_0 \sim q(x)$, noise is continuously added to the distribution, and the standard deviation of the noise is a fixed value, and the mean is determined by the fixed value and the current data at time t. This process comes from a Markov chain, and as t increases, the final data distribution X_T becomes an isotropic Gaussian distribution. The reconstruction process is an inverse diffusion process, recovering original data from Gaussian noise. However, it is unable to fit the distribution gradually, requiring the construction of a parameter distribution for estimation. The reconstruction process remains a Markov chain, where a U-Net network [15] can be trained during the reconstruction process to predict the denoising process. As shown in Figure 2, noise is added to the input image x_0, which can be completed in one step, that is, T steps of noise addition are completed once. Then, a U-Net network is trained to pre-
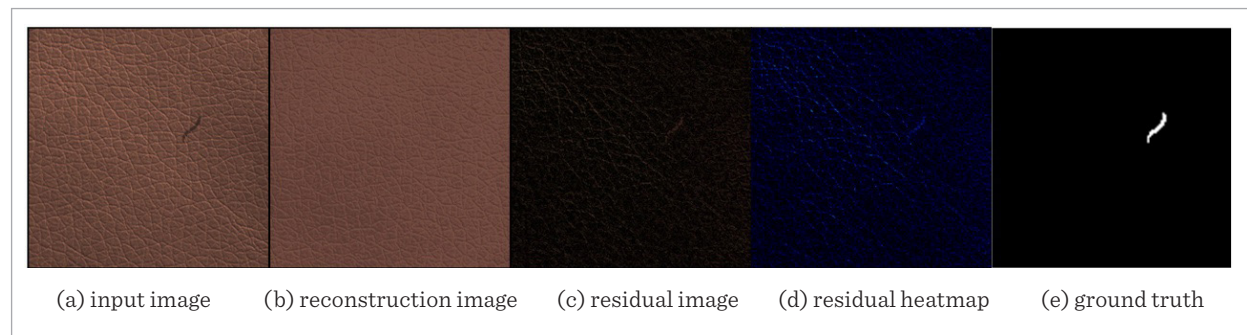
Figure 2
DDPM Generator
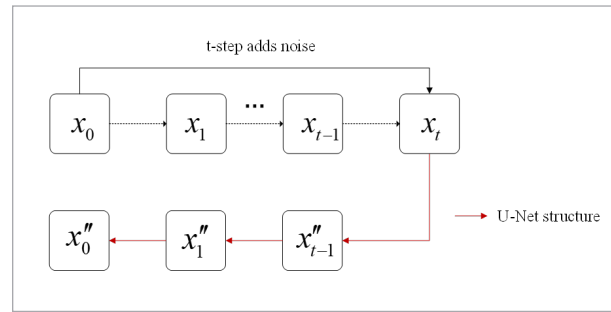
dict the noisy image. In each step of the reconstruction process, the trained U-Net network is used for denoising, aiming to reconstruct the original image x_0'' as much as possible. During inference, defective images can be reconstructed into defect-free images.

The full-length Markov chain diffusion process increases the time cost of training. This paper proposes an anomaly detection method based on reconstruction, which does not require a full-length Markov chain. By parameterizing noisy steps to time t, the noisy image x_t is obtained, and the reconstruction process from x_t is started, ultimately the reconstructed image x_0'' is generated. At the training stage of the DDPM generator, train on normal samples, and use partially long Markov chains for the diffusion process. In the inference stage, reconstruct images with added noise, and abnormal samples to be reconstructed into normal images, normal samples to be reconstructed into their original images. Figure 3 shows the reconstruction results of the DDPM generator for the defective leather dataset. Figure 3(a) represents the input defective original image, Figure 3(b) is the

Figure 3
Reconstruction result of DDPM generator on leather image

(a) input image   (b) reconstruction image   (c) residual image   (d) residual heatmap   (e) ground truth

reconstructed image of DDPM generator for Figure 3(a). Figure 3(c) is the residual image of Figure 3(a) and Figure (b). Figure 3(d) is the residual heatmap, and Figure 3(e) is the ground truth. The defective parts in the image can be observed in Figures 3(c)-(d), but they are not very prominent.
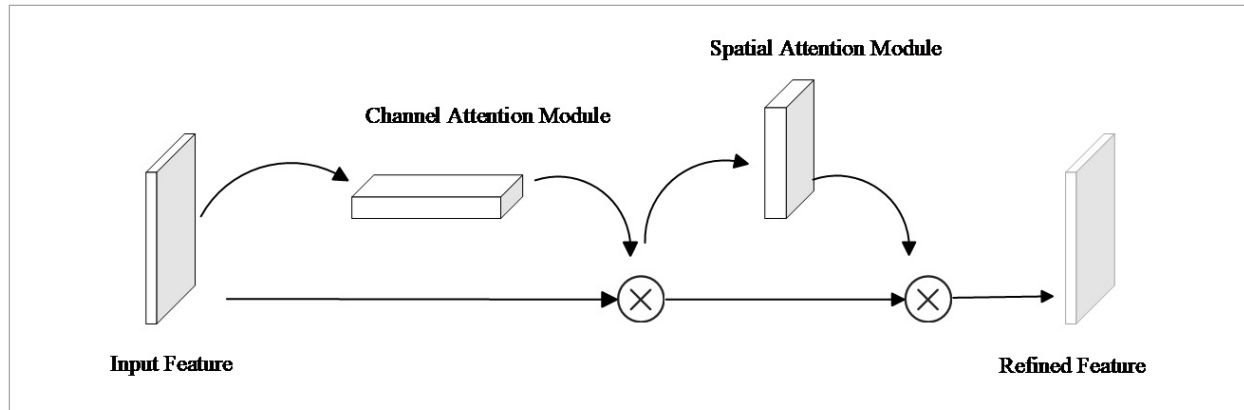
2.3. AE Generator

When reconstructing abnormal image samples using Autoencoders (AE), it is prone to reconstructing the easily abnormal parts. To enhance AE's reconstruction capability, a CBAM module [21], as shown in Figure 4, was incorporated in the encoder.

Figure 4
CBAM module added to the encoder

The CBAM module combines channel attention and spatial attention, enabling the network to focus more on normal image information through the CBAM module. In the inference stage, the image with defective parts can be better reconstructed into a normal image. The decoder can reduce the difficulty of reconstruction through Skip Connection; however, if Skip Connection is added to each layer, it will reconstruct all defects. To address the issue of all defects being reconstructed, this study only adds Skip Connection to the bottom layer, thereby making it less likely for defects to be easily reconstructed during the reconstruction process.

2.4. Feature Fusion

The network structure proposed in this paper includes two generators. During inference, the reconstructed results of the two generators need to be feature fused and then input into the discriminator for discrimination. Since the semantic information of the corresponding channels in the generated results of the two generators is similar, direct feature addition is chosen for fusion, which can retain more information of the generated images and then input them into the discriminator.

2.5. Training

In this paper, a DDPM generator is trained on normal data samples. During the training, the input images are all normal images, and partial Markov chain is used to denoise the images during the forward diffusion process $q\left(x|x_{t-1}\right)$,

$$q(x_t \,|\, x_{t-1}) \,=\, N\left(x_t \,|\, x_{t-1}\sqrt{1-\beta_t}\,, \beta_t\mathrm{I}\right) \qquad (1)$$

for $t = 1,\ldots,T$. The noise process is regulated by variance $\beta_t \in \left(0,1\right), t = 1,\ldots,T$, and we can add noise by defining a linear schedule, for β_1 from 10^{-4} to 0.02. The reconstruction process takes θ as a parameter and takes noise-added images as input and samples according to the following formula:

$$P_\theta\left(x_{t-1} \,|\, x_t\right) = N\left(x_{t-1} \,|\, \mu_\theta\left(x_t \,|\, t\right), \beta_t I\right) \qquad (2)$$

for $t = T,\ldots,1$, $\alpha_t = 1 - \beta_t$, $\overline{\alpha}_t = \prod_{i=0}^{T}\alpha_i$, $\beta_t = \dfrac{1-\overline{\alpha}_{t-1}}{1-\overline{\alpha}_t}\beta_t$,

μ_θ can be used to implement by a structure similar to U-Net, make the DDPM generator can learn how to put a image of a containing noise denoising to normal image. Objective function Using simplified objective function L_s [8]:

$$L_s = E_{t \sim [1-T], x_0 \sim q(x_0), \varepsilon \sim N(0,I)} \left[\left\| \varepsilon - \varepsilon_\theta \left(x_t, x \right) \right\|^2 \right] \quad (3)$$

the parameters ε_θ can be learned through the network.

When training AE generator, it is also trained on the normal data samples, followed by adversarial training with the discriminator. Both training processes are unsupervised. Due to the contribution of adversarial loss, context loss, and latent loss to the training objective, the three loss values are weighted and combined during training in this paper.

In order to achieve the best reconstruction capability, the adversarial loss mentioned by Goodfellow et al. [7] is utilized to ensure that the generative network G can reconstruct the input image x as close to the real sample $\hat{x}$ as possible. Moreover, the discriminator D is trained to distinguish between real samples and fake generated samples. This loss is denoted as L_{adv}, as shown in Formula (4):

$$L_{adv} = E_{x \sim p_x} \left[log \, D(x) \right] + E_{x \sim p_x} \left[log \left(1 - D(\hat{x}) \right) \right] \quad (4)$$

In order to learn the context information of normal samples, the input image and reconstructed image are L_1 normalized to ensure that the reconstructed graph is similar to the context information of normal samples This loss is represented as L_{con}:

$$L_{con} = E_{x \to p_x} |x - \hat{x}|_1 \quad (5)$$

L_{adv} and L_{con} enables the model to generate realistic and context-similar images, in order to enable the model to reconstruct the potential distribution of the normal image from the input image x to be consistent, this paper uses the last convolution layer of the discriminator D to extract x and $\hat{x}$ features to reconstruct its potential representation. The loss is denoted as L_{lat}:

$$L_{lat} = E_{x \to p_x} |f(x) - f(\hat{x})|_2. \quad (6)$$

The weighted sum of the three losses is taken as the total training objective L, as shown in formula (7), where, λ_{adv}, λ_{con} and λ_{lat} are the weight parameters:

$$L = \lambda_{adv} L_{adv} + \lambda_{con} L_{con} + \lambda_{lat} L_{lat}. \quad (7)$$

2.6. Inference

In inference stage, image is input into the DDPM generator and the AE generator. After entering the DDPM generator, part of the Markov chain is used for noise addition during the diffusion process. In the reconstruction process, the denoised reconstruction is performed on the noisy image, allowing the original image to be reconstructed with the same effect. Defective images are reconstructed to be defect-free. Similarly, when an image enters the AE generator, a reconstructed image is obtained. Two reconstructed images are obtained from the two generators. These two images are fused by feature-wise addition and input into the discriminator for image identification by calculating the difference between the latent representation and the difference between the reconstructed image and the original image.

3. Experiments

3.1. Dataset

The network proposed in this paper has been experimented and evaluated on the MVTec AD dataset [3], which contains images of industrial production, including 5334 high-resolution color images, five categories of texture images from different fields, and ten categories of object structure images. Each category consists of a training set and a testing set. The training set only contains normal samples, while the testing set contains both normal and defective samples. The resolution of the images is between 700 and 1024, and the resolution of the images is reduced to 256 in the experiment. Figure 5 shows an abnormal sample of each category of images. The first row represents part of the abnormal texture image samples, and the second and third rows represent part of the abnormal object structure image samples.

3.2. Implementation Details

The AE generator includes an encoder and a decoder. In the encoder structure, this paper adds CBAM attention module to improve the performance of the network, so that the network can pay more attention to the normal image information. To retain global and local image information, the Skip Connection structure in U-net network is introduced to enhance the

Figure 5
Display of defect images in MVTec AD dataset
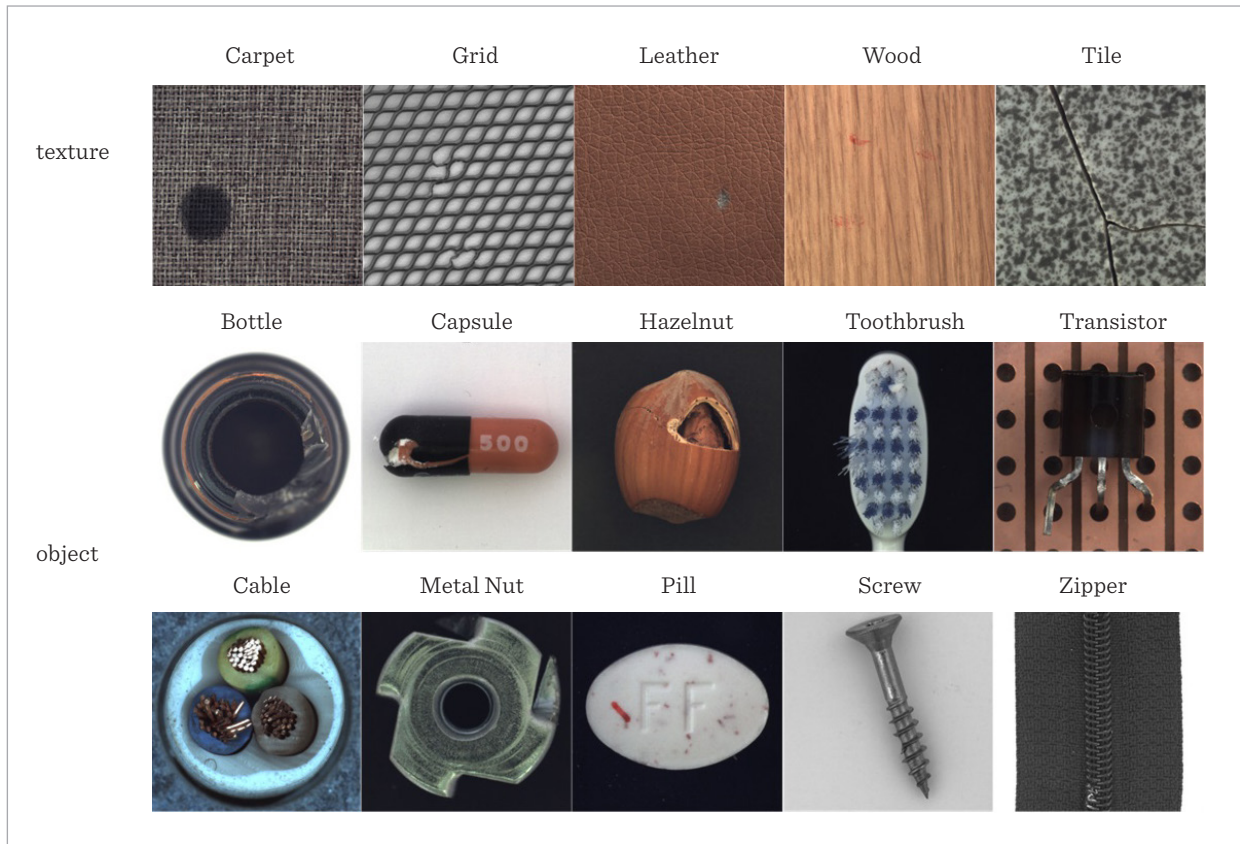
network's reconstruction capability, but to avoid the network to reconstruct more defects, the Skip Connection structure is only added in the last layer of the network. During training, set epoch as 50 and the batch size as 100, use Adam optimizer [10] to optimize the target L, set the learning rate as $2e^{-3}$, use lambda decay, $\beta_1 = 0.5$, $\beta_2 = 0.999$. The weighted parameters λ_{adv}, λ_{con}, λ_{lat} of L are selected as 1, 40 and 1 respectively after the experiment.

During the training of DDPM generator, use the improved U-Net in Dhariwal et al. [4] to train the de-noising model, and use the Adam optimizer for training. The learning rate was set as $1e^{-4}$, the batch size was 32, and the epochs was 100. Considering that full-length Markov chain diffusion is not required for anomaly detection based on reconstruction, the number of steps is parameterized and the partial length Markov chain is used for diffusion.

3.3. Ablation Study

To evaluate whether the improvement in this paper enhances the model's performance, ablation experiments were conducted on the model. The ablation experiments include comparing the SkipGANomaly algorithm with no attention mechanism and full Skip Connection in the AE generator (Algorithm 1), the AE generator algorithm with attention mechanism in the encoder and only one layer of Skip Connection (Algorithm 2), and the proposed DGRNet with dual generators based on Algorithm 2, adding a DDPM generator. The results are shown in Tables 1-2; Table 1 shows the results for texture images and Table 2 for object images.

The evaluation metric used is the area under the ROC curve (AUC). From Table 1, it can be seen the AUC for texture images, and from Table 2, it can be seen the AUC for object structure images. From these two tables, it can be seen the improvement of the AE gener-

Table 1

Comparison of AUC results of texture images in ablation experiments

Category	Algorithm 1	Algorithm 2	DGRNet
Carpet	41	42.3	**45.2**
Grid	47.7	**76.9**	69.3
Leather	60.9	**65.3**	64.2
Tile	29.9	**74.6**	72.1
Wood	19.9	**57.5**	56.9
avg. textures	39.5	**63.3**	61.5

Table 2

Comparison of AUC results of object structure images in ablation experiments

Category	Algorithm 1	Algorithm 2	DGRNet
Bottle	52.7	**46.4**	**56.9**
Cable	48.1	**64**	55.2
Capsule	50	57	**61.9**
Hazelnut	24.5	**64.9**	51.9
Metal Nut	33.7	51.4	**52.1**
Pill	34.8	40	**51.5**
Screw	46.4	18.3	**89.9**
Toothbrush	55.6	42.5	**58.6**
Transistor	48.5	55.6	**61.8**
Zipper	47.6	**53.3**	51.9
avg. objects	44.2	49.3	**59.2**

ator significantly outperforms the baseline network SkipGANomaly algorithm. The detection results on 12 category datasets have been improved. The DGRNet in this paper outperforms Algorithm 2, especially showing significant improvement on the dataset of object structure type, demonstrating the strong generation and reconstruction capabilities of the DDPM generator for image of object structure type. The average AUC of the three algorithms for the MVTecAD data set is 42.6, 55, 60 respectively. It can be seen from the results of the ablation experiment that the improvement of the network structure in this paper is very useful to improve the performance of the network.

3.4. Experiments Results

The model was evaluated using AUC, which is a performance metric for assessing anomaly detection, the value of AUC ranges from 0.1 to 1.0, A higher value indicates better model performance. In this study, we compared DGRNet with AnoGAN, GANomaly and SkipGANomaly models. The results are shown in Tables 3-4. We divided the dataset into two categories: texture and object structure, Table 3 is the results for texture images and Table 4 is for object structure images. On average, the model proposed in this study outperforms the other network models in both texture and object structure datasets. The average AUC of the AnoGAN, GANomaly, SkipGANomaly and DGRNet models for the MVTecAD data set is 40.4, 42.1, 42.6 and 60, respectively. It can be seen that DGRNet in this paper is 19.6 percentage points higher than the baseline network. Furthermore, the DGRNet model exhibits more stable performance across the fifteen category datasets.

Table 3

AUC results of GANomaly, SkipGANomaly, and the model DGRNet in this paper of texture images in MVTec AD dataset

Category	AnoGAN[19]	GANomaly[1]	SkipGANomaly[2]	DGRNet
Carpet	20.4	20.3	41	**45.2**
Grid	22.6	40.4	47.7	**69.3**
Leather	37.8	41.3	60.9	**64.2**
Tile	17.7	40.8	29.9	**72.1**
Wood	38.6	**74.4**	19.9	56.9
avg. textures	27.4	43.4	39.5	**61.5**

Table4
AUC results of GANomaly, SkipGANomaly, and the model DGRNet in this paper of object images in MVTec AD dataset

| Category	AnoGAN[19]	GANomaly[1]	SkipGANomaly[2]	DGRNet
Bottle	62	25.1	52.7	**56.9**
Cable	38.3	45.7	48.1	**55.2**
Capsule	30.6	**68.2**	50	61.9
Hazelnut	50	**53.7**	24.5	51.9
Metal Nut	32	27	33.7	**52.1**
Pill	**62**	47.2	34.8	51.5
Screw	35	23.1	46.4	**89.9**
Toothbrush	57	37.2	55.6	**58.6**
Transistor	54.9	44	48.5	**61.8**
Zipper	46.7	43.4	47.6	**51.9**
avg. objects	46.9	41.5	44.2	**59.2**

Taking the leather dataset as an example, Figure 6 displays the reconstruction results of DGRNet model on the leather dataset. Figure 6(a) represents the input defective original image, Figure 6(b) is the reconstructed image for Figure 6(a), Figure 6(c) is the residual image of Figures 6(a)-(b). Figure 6(d) is the residual heatmap, and Figure 6(e) is the ground truth. The defective parts in the image can be observed in Figures 6(c)-(d). Compared Figure 3 that with the reconstruction results obtained only using the DDPM generator mentioned above. It can be observed that the detection effect of the proposed network is more significant, with the defective parts being reconstructed closer to normal images. The residual heatmap also more clearly reveals the previously defective abnormal regions. When compared with the true labels, the results are closer to the true labels, indicating that the proposed method can effectively reconstruct the image more efficiently.

Figure 6
Reconstruction results of leather dataset through DGRNet model

(a) input image   (b) reconstruction image   (c) residual image   (d) residual heatmap   (e) ground truth
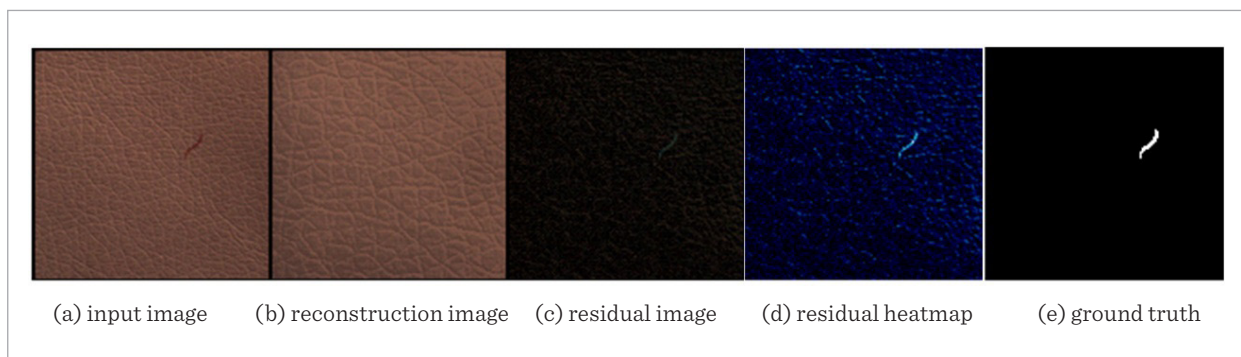
4. Conclusion

The proposed dual-generator reconstruction network in this paper is a novel algorithm based on the reconstruction network and has been experimented on industrial images. This algorithm is an unsupervised method consisting of two generators and a discriminator. One of the generators utilizes the powerful generation capability of the DDPM to enhance the reconstruction ability of the reconstruction network, while the other generator is an AE with Skip Connection and attention mechanisms to improve its generation capability. The DDPM has recently been garnering attention but has not been widely applied in the field of industrial image anomaly detection. This network applies DDPM to the field of anomaly detection, leveraging its powerful generation capability. In this paper, two generators are used to reconstruct the input image and the reconstructed images of the two generators are fused to improve the generating ability of the network. The algorithm has been verified on the MVTec AD dataset, showing significant improvements compared to the baseline network's detection results on the same dataset.

Due to the lack of computing power in the experimental environment, the number of iterations and the training batch size set in the experiment of training DDPM generator are small, so the generation ability of DDPM may not be fully developed, and there is still a certain gap between the results of the best industrial image anomaly detection algorithms. In the following work, DDPM generator will be further improved to give full play to its generation ability and improve its training speed. Relevant experiments also have done on other data sets, such as breast ultrasound data set, but the results are not very satisfactory, which may be related to the low quality of ultrasound images themselves. After that, image preprocessing and network structure optimization may improve the detection performance on this kind of data set. On the data set of fabric texture class, the performance is very good, and it can be applied to fabric defect detection in the future, but the detection effect on object structure images is significantly improved. The next step will also be to improve this problem.

Acknowledgement

References

1. Akcay, S., Atapour-Abarghouei, A., Breckon, T. P. GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training. Computer Vision - ACCV 2018, 2019, 622-637. https://doi.org/10.1007/978-3-030-20893-6_39

2. Akcay, S., Atapour-Abarghouei, A., Breckon, T. P. Skip-GANomaly: Skip Connected and Adversarially Trained Encoder-Decoder Anomaly Detection. 2019 International Joint Conference on Neural Networks (IJCNN), 2019. https://doi.org/10.1109/IJCNN.2019.8851808

3. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C. MVTec AD - A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, 9584-9592. https://doi.org/10.1109/CVPR.2019.00982

4. Dhariwal, P., Nichol, A. Diffusion Models Beat GANs on Image Synthesis. ArXiv:2105.05233 [Cs, Stat]. 2021. https://arxiv.org/abs/2105.05233

5. Du, X., Li, B., Zhao, Z., Jiang, B., Shi, Y., Jin, L., Jin, X. Anomaly-Prior Guided Inpainting for Industrial Visual Anomaly Detection. Optics & Laser Technology, 2024, 170, 110296. https://doi.org/10.1016/j.optlastec.2023.110296

6. Farady, I., Kuo, C.C., Ng, H.F., Lin, C.Y. Hierarchical Image Transformation and Multi-Level Features for Anomaly Defect Detection. Sensors, 2023, 23(2), 988. https://doi.org/10.3390/s23020988

7. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. Generative Adversarial Networks. Communications of the ACM, 2020, 63(11), 139-144. https://doi.org/10.1145/3422622

8. Ho, J., Jain, A., Abbeel, P. Denoising Diffusion Probabilistic Models. ArXiv:2006.11239 [Cs, Stat]. 2020. https://arxiv.org/abs/2006.11239

9. Hodge, V., Austin, J. A Survey of Outlier Detection Methodologies. Artificial Intelligence Review, 2004, 22(2), 85-126. https://doi.org/10.1023/B:AIRE.0000045502.10941.a9

10. Kingma, D. P., Ba, J. Adam: A Method for Stochastic Optimization. ArXiv.org, 2014. https://arxiv.org/abs/1412.6980

11. Lv, C. K., Shen, F., Zhang, Z. T., Zhang, F. Review of Image Anomaly Detection. Acta Automatica Sinica, 2022, 48(6), 1402–1428.

12. Mei, S., Yang, H., Yin, Z. An Unsupervised-Learning-Based Approach for Automated Defect Inspection on Textured Surfaces. IEEE Transactions on Instrumentation and Measurement, 20181, 67(6), 1266-1277. https://doi.org/10.1109/TIM.2018.2795178

13. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M. Hierarchical Text-Conditional Image Generation with CLIP Latents. ArXiv:2204.06125 [Cs]. 2022. https://arxiv.org/abs/2204.06125

14. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. ArXiv:2112.10752 [Cs]. 2022. https://arxiv.org/abs/2112.10752 https://doi.org/10.1109/CVPR52688.2022.01042

15. Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Lecture Notes in Computer Science, 2015, 9351, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

16. Rumelhart, D. E., Hinton, G. E., Williams, R. J. Learning Representations by Back-Propagating Errors. Nature, 1986, 323(6088), 533-536. https://doi.org/10.1038/323533a0

17. Sabokrou, M., Khalooei, M., Fathy, M., Ehsan Adeli. Adversarially Learned One-Class Classifier for Novelty Detection. ArXiv (Cornell University). 2018, https://doi.org/10.1109/CVPR.2018.00356

18. Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S. K. S., Ayan, B. K., Mahdavi, S., Lopes, R. G., Salimans, T., Ho, J., Fleet, D. J., Norouzi, M. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. ArXiv:2205.11487 [Cs]. 2022. https://arxiv.org/abs/2205.11487 https://doi.org/10.1145/3528233.3530757

19. Schlegl, T., Seeböck, P., Waldstein, S. M., Schmidt-Erfurth, U., Langs, G. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. Lecture Notes in Computer Science, 2017, 146-157. https://doi.org/10.1007/978-3-319-59050-9_12

20. Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., Ganguli, S. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. ArXiv:1503.03585 [Cond-Mat, Q-Bio, Stat]. 2015. https://arxiv.org/abs/1503.03585

21. Woo, S., Park, J., Lee, J.-Y., Kweon, I. S. CBAM: Convolutional Block Attention Module. Computer Vision - ECCV 2018, 2018, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1

22. Wyatt, J., Leach, A., Schmon, S. M., Willcocks, C. G. AnoDDPM: Anomaly Detection with Denoising Diffusion Probabilistic Models using Simplex Noise | IEEE Conference Publication | IEEE Xplore. 022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022, 649-655. https://doi.org/10.1109/CVPRW56347.2022.00080

23. Zhao, J., Mathieu, M., Yann LeCun. Energy-based Generative Adversarial Network. ArXiv (Cornell University). 2016.

24. Zhou, K., Li, J., Xiao, Y., Yang, J., Cheng, J., Liu, W., Luo, W., Liu, J., Gao, S. Memorizing Structure-Texture Correspondence for Image Anomaly Detection. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(6), 2335-2349. https://doi.org/10.1109/TNNLS.2021.3101403