International Journal of Systems, Control and Communications

ISSN online: 1755-9359 - ISSN print: 1755-9340
https://www.inderscience.com/ijscc

Skip.Self attention GAN for anomaly detection

Hanlin Sun, Yefeng Liu, Yihang Ma, ChongZheng Na, Qichun Zhang

DOI: 10.1504/IJSCC.2025.10069429

Article History:

Received:	02 October 2024
Last revised:	07 November 2024
Accepted:	18 November 2024
Published online:	18 February 2025

Skip.Self attention GAN for anomaly detection

Hanlin Sun

School of Automation and Electrical Engineering,
Shenyang Ligong University,
Shenyang 110159, China
Email: 836827498@qq.com

Yefeng Liu*

School of Automation and Electrical Engineering,
Shenyang Ligong University,
Shenyang 110159, China
and
Liaoning Key Laboratory of CNC Machine Tool
Cyber-Physical Fusion and Intelligent Manufacturing,
Shenyang Institute of Technology,
Shenfu Demonstration Zone, Liaoning 113122, China
Email: liuyefeng@situ.edu.cn
*Corresponding author

Yihang Ma and ChongZheng Na

Liaoning Key Laboratory of CNC Machine Tool
Cyber-Physical Fusion and Intelligent Manufacturing,
Shenyang Institute of Technology,
Shenfu Demonstration Zone, Liaoning 113122, China
Email: mayihang@situ.edu.cn
Email: nachongzheng@situ.edu.cn

Qichun Zhang

School of Creative and Digital Industries,
Buckinghamshire New University,
High Wycombe, HP11 2JZ, UK
Email: q.zhang17@bradford.ac.uk

Abstract: Generative adversarial network (GAN) has low discrimination ability for abnormal samples or noise data in anomaly detection tasks, which affects the detection effect. This paper proposes an improved model based on Skip-GANomaly network, which adds a self-attention module to the encoder of the generator to improve the model's ability to capture long-distance dependence of input data, so as to better understand the overall structure and local details of the image. Moreover, it dynamically focuses on the most important parts of the anomaly detection task, such as the shape, texture and colour of the object, so that the data representation ability and anomaly detection accuracy of the model are significantly improved. MvTec public dataset and KolektorSDD dataset were used for experimental verification. The results show that compared with AnoGAN, GANomaly, and Skip-GANomaly, it has better performance in terms of receiver operating characteristic (ROC) area under the curve (AUC), which proves the effectiveness of this model and has a certain application prospect.

Keywords: generative adversarial network; GAN; anomaly detection; Skip-GANomaly; self-attention.

Reference to this paper should be made as follows: Sun, H., Liu, Y., Ma, Y., Na, CZ. and Zhang, Q. (2025) 'Skip.Self attention GAN for anomaly detection', *Int. J. Systems, Control and Communications*, Vol. 16, No. 1, pp.17–32.

Biographical notes: Hanlin Sun is a Master's student and his research areas include deep learning and computer vision.

Yefeng Liu is a Professor in Liaoning Key Laboratory of Information Physics Fusion and Intelligent Manufacturing for CNC Machine, Shenyang Institute of technology, Shenfu New District, China. His current research interests include the development of manufacturing execution systems, production planning and scheduling, and intelligent optimisation methods.

Yihang Ma is a teacher in Liaoning Key Laboratory of CNC Machine Tool Cyber-Physical Fusion and Intelligent Manufacturing, Shenyang Institute of Technology, Shenfu Demonstration Zone, Liaoning 113122, China.

ChongZheng Na is a teacher in Liaoning Key Laboratory of CNC Machine Tool Cyber-Physical Fusion and Intelligent Manufacturing, Shenyang Institute of Technology, Shenfu Demonstration Zone, Liaoning 113122, China.

Qichun Zhang is a teacher in School of Creative and Digital Industries, Buckinghamshire New University, High Wycombe, HP11 2JZ, UK.

1 Introduction

For different industrial products, quality issues are the most important (Chen, 2004). The income and reputation of the business will suffer if the provided products have flaws. The most significant issue is that users might be impacted by the functional flaws in the product, which could lead to security issues or even catastrophic mishaps. Consequently, in order for businesses to create goods at this time, fault detection of industrial items is a prerequisite (Zhou et al., 2017). Using the manual detection line, traditional defect identification is primarily done by eye (Zhang et al., 2020). There will be issues with low efficiency and a high missed detection rate, nevertheless, as product production increases quickly and precision continues to be improved. Low-cost, high-efficiency techniques must be developed based on conventional anomaly detection methods (Luo et al., 2022), and machine vision-based deep learning techniques must be gradually implemented in (Hu et al., 2020). In this instance, the deep learning model-based defect detection technique can assist the system in capturing the finer details and flaws of the industrial product surface, enhancing the precision and effectiveness of defect identification. The use of this technology is anticipated to have a significant impact on increasing production effectiveness and product quality, as well as opening up new avenues for industrial product quality inspection.

Traditional object identification methods, on the other hand, require more supervised learning, which raises the cost because it requires gathering enough samples for labelling and classification (Lin et al., 2021). Furthermore, the majority of photos that can be gathered for real businesses are normal, defect-free samples, with a very tiny percentage of defective samples. This will restrict the training and subsequent detection of standard detection algorithms. Unsupervised learning techniques based on deep learning have increasingly gained popularity in recent years as a solution to this issue. These techniques may be more advantageous for few-shot defect identification issues because they do not require a lot of labelled data and can develop feature representations from unlabeled data. Goodfellow et al. (2020) proposed generative adversarial network (GAN) for the first time, and since then, GAN network and its derivative networks have been applied to various fields. Schlegl et al. (2017) firstly applied the generative adversarial network to anomaly detection, and proposed anomaly detection generative adversarial network (AnoGAN). It only learned how positive instances were distributed, and the input image's feature distribution and the reconstructed image's feature distribution ought to be nearly identical. Then the abnormal image is determined by the residual between them. Later, Schlegl et al. (2019) proposed fast unsupervised anomaly detection with generative adversarial networks (F-AnoGAN) in 2019, which can learn latent space features more quickly. Zenati et al. (2018b) proposed efficient GAN-based anomaly detection (EGBAD), which shows the advantages of the joint learning of latent representation and reconstruction map. GANomaly is an adversarial autoencoder proposed by Akcay et al. (2018) to further strengthen the reconstruction ability of the model, which detects anomalies by comparing the latent representation of the input image with that of the generated image. After that, skip connected and adversarially trained encoder-decoder anomaly detection (Skip-GANomaly) was also proposed (Akay et al., 2019). Is a method that adds U-net network (Ronneberger et al., 2015) to GANomaly generator and improves the quality of image reconstruction from latent space through skip connection. Skip-attention generative adversarial networks (SAGAN) (Liu et al., 2021) is a generative adversarial network combined with attention mechanism, which has achieved significant performance improvement in industrial product anomaly detection tasks. By introducing the attention mechanism, SAGAN can better capture the correlation and important information of different regions in the image, so as to effectively improve the detection accuracy and robustness of industrial product defects. Combining the benefits of generative adversarial networks, it allows the model to automatically identify and extract the important features from the image through the discriminator and generator's adversarial learning. This allows for the effective detection and recognition of industrial product defects. Skip-GANomaly is a cutting-edge technology that has garnered a lot of attention in the field of industrial product quality detection because of its advantages, which include great resilience, flexibility and adaptability, feature extraction combined with skip connection, and generalisation. However, the model might not successfully preserve the intricacies in the original input data during the encoding and decoding process because of the focus on high-level feature representation.

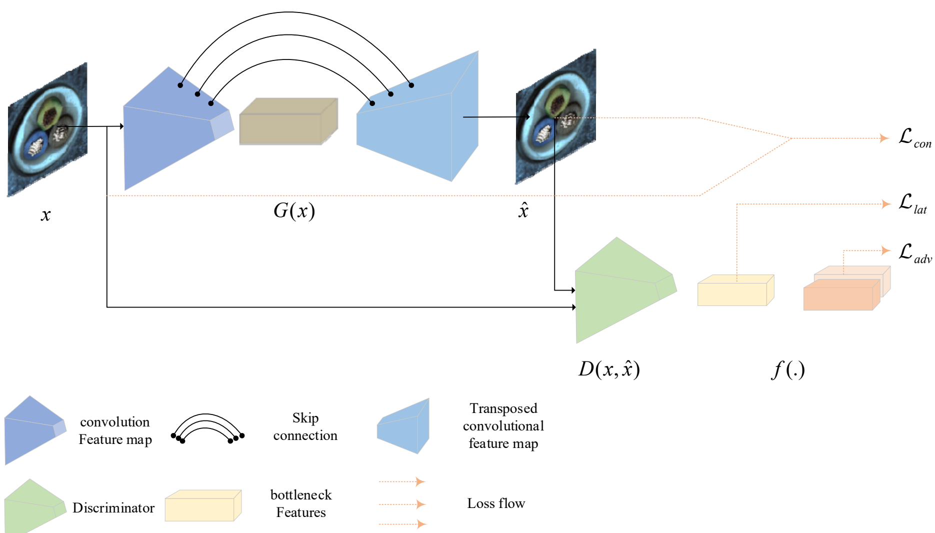
In order to solve the problem of missed detection and error detection caused by the strong reconstruction ability of Skip-GANomaly, the self-attention generative adeversarial networks (Zhang et al., 2018) was combined with Skip-GANomaly. A self-attention mechanism is added to the U-net network. In order to adapt to the requirements of real scenarios, MvTec (Bergmann et al., 2020) and KolektorSDD (Tabernik et al., 2020) were used as datasets for experiments. The advantages of this model are proved by comparative tests.

The following is the order of the article. The internal construction and general structure of Skip-GANomaly are introduced in Section 2. The details of the updated network structure are presented in Section 3. The datasets used are introduced in Section 4, and the experimental findings are displayed in Section 5. The conclusion is covered in Section 6.

2 Skip-GANomaly model

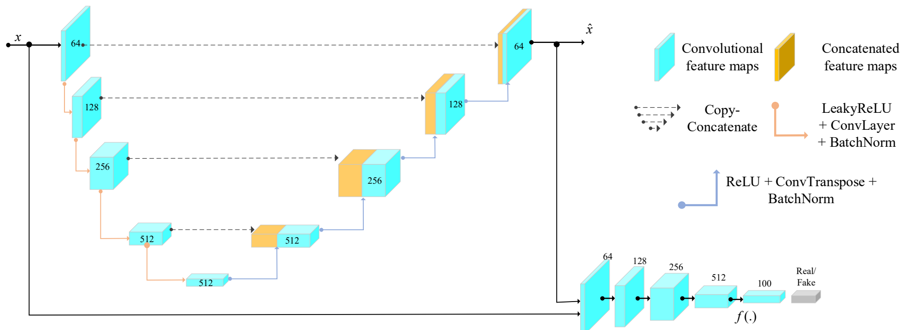
For the Skip-GANomaly model, Figure 1 consists of the generator (G) and discriminator (D) networks. The generator (G) consists of encoder (G_E) and decoder (G_D) to form a 'bow-tie' U-net convolutional neural network. The encoder (G_E) captures the feature distribution of the input by mapping a high-dimensional image x into a low-dimensional latent space z, let G_E: $G_E: x \to z$, where $x \in R^{w \times h \times c}$, $z \in R^d$.

Figure 1 Skip-GANomaly model (see online version for colours)



In Figure 2, the network GE will process the input x through five blocks containing the convolutional layer, BatchNorm layer and LeakyReLU activation function to generate the latent representation z, which is the bottleneck feature. This feature will uniquely express the input information. The decoder network GD maps the latent vector z back to the dimensions of the input image and reconstructs the output, denoted as $\hat{x}$. This symmetry is similar to GE and ensures that the decoder can reproduce the image efficiently. According to the method in Ronneberger et al. (2015), the decoder GD is designed by using the skip connection method, such an operation can make each individual downsampling layer in the network be connected to the corresponding up-sampling layer. By skipping connections, information can be passed directly between layers, providing a significant advantage. This approach preserves both local and global (multi-scale) information and thus leads to better reconstruction results.

Figure 2 Internal structure diagram of the generator G_E (see online version for colours)



As a discriminator network, its role is to predict and discriminate the input class, and classify the fake image $\hat{x}$ generated by the generator G and the real image x. Its network architecture is the same as DCGAN (Radford et al., 2015). The discriminator D is also used as a feature extractor to compute the latent representation between the reconstructed image $\hat{x}$ and the real image x.

The loss function of the network is divided into adversarial loss (Goodfellow et al., 2020), context loss (Gatys et al., 2018) and potential loss (Sun et al., 2021).

1 Adversarial loss: as shown in equation (1), this loss keeps the generator G to generate more realistic image $\hat{x}$ as much as possible, and the discriminator D to classify real samples and fake samples (generated pictures). The function works the same way as in the GAN network, minimising the objective of the generator G and maximising the objective of the discriminator D.

$$l_{adv} = \underset{x \sim p_x}{E}[\log D(x)] + \underset{x \sim p_x}{E}\left[\log\left(1 - D(\hat{x})\right)\right] \tag{1}$$

2 Context loss: as shown in equation (2), in order to fully learn the context information and better capture the input data distribution of positive samples, L_1 normalisation is applied to the input x and output $\hat{x}$ to ensure that the model can generate images with high contextual similarity to normal samples.

$$l_{con} = \underset{x \sim p_x}{E}|x - \hat{x}|_1 \tag{2}$$

3 Potential loss: as shown in equation (3), in order to be able to reconstruct the latent representation of the input x and make the generated sample $\hat{x}$ as close as possible to the input, the last convolution layer in the discriminator D is used to extract the features of x and $\hat{x}$ and reconstruct their latent representations, let $z = f(x)$, $\hat{z} = f(\hat{x})$.

$$l_{lat} = \underset{x \sim p_x}{E}\left|f(x) - f(\hat{x})\right| \tag{3}$$

3 Improve the Skip-GANomaly network model

Skip-GANomaly is a cutting-edge technology that has garnered a lot of attention in the field of industrial product quality detection because of its advantages, which include great resilience, flexibility and adaptability, feature extraction combined with skip connection, and generalisation. However, the model might not successfully preserve the intricacies in the original input data during the encoding and decoding process because of the focus on high-level feature representation.

In order to solve the problem of missed detection and error detection caused by the strong reconstruction ability of Skip-GANomaly, the self-attention generative adeversarial networks was combined with Skip-GANomaly, and proposed an improved Skip-GANomaly model.

3.1 Self-attention mechanism

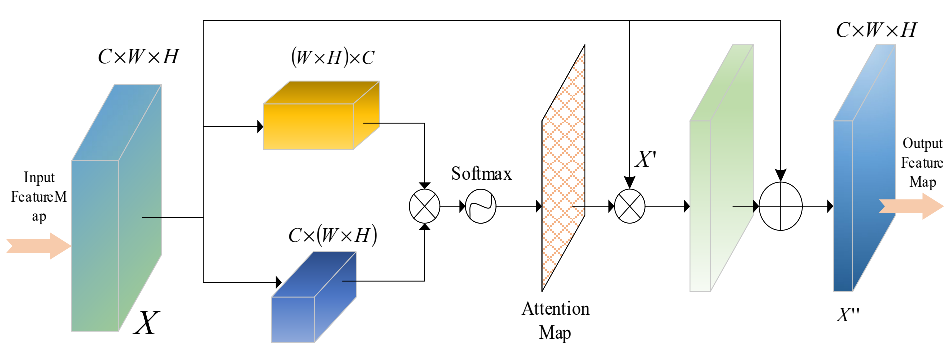
Self-attention is an integrated mechanism in deep learning models, which allows the model to take into account the entire input sequence information when processing a given element. In the field of image processing, this means that the network can consider the information of each pixel in the image at the same time. The self-attention mechanism is particularly suitable for capturing long-distance dependencies between data, which is difficult to implement in traditional convolutional neural networks.

After fusing self-attention into each convolutional layer of the encoder in this model, the input feature map is fed into the subsequent self-attention module. As shown in Figure 3, given the feature map $X \in R^{C \times H \times W}$, the attention map $M_a \in R^{N \times N}$ is first generated by calculating the relationship between the feature maps, where $N = H \times W$. This feature is then sent to two different 1×1 convolutional layers to generate two new feature maps A, $B \in R^{C \times H \times W}$, which are then reshaped into $R^{C \times N}$ and matrix multiplications are performed on A and B^T, respectively. Finally, softmax is used to normalise the weights. The calculation formula of feature map is equation (4):

$$M_a = \sigma\left(A^T \cdot B\right) \tag{4}$$

where σ represents the softmax function.

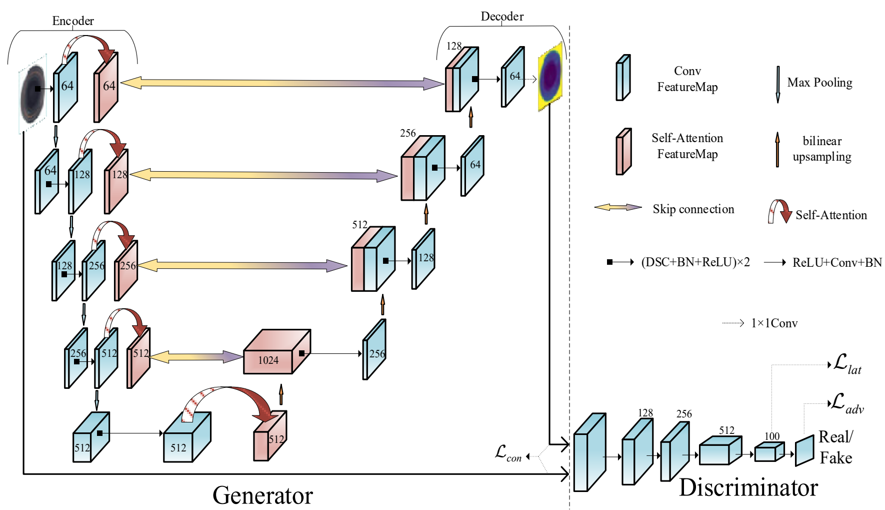
Figure 3 Structure of self-attention (see online version for colours)



3.2 Skip-GANomaly network improvement

The Skip-GANomaly network was used as the basic architecture, and the network was modified to solve the shortcomings of its detection process. The network architecture of this model is shown in Figure 4, which is composed of a generator *G* and a discriminator *D*. The generator, like the original model, uses a U-net convolutional neural network to ensure the quality of the generated image. The difference is the introduction of self-attention mechanism in the generator, which uses global pooling instead of convolution in the downsampling process and bilinear interpolation instead of transposed convolution in the upsampling process. At the same time, the same depthwise separable convolution (DSC) as SAGAN was used to reduce the number of model parameters, so as to effectively reduce the computational cost and improve the training speed of the model.

Figure 4 Improved skip-GANomaly model (see online version for colours)



In Figure 4, it can be seen that the generator part consists of a U-shaped encoder-decoder structure. In the encoder, the feature map of each layer first goes through two DSC to increase the number of channels, and then the feature map is extracted through the self-attention mechanism. We then use max pooling to reduce its size by half. The self-attention feature map of each layer is skip connected with the feature map of the corresponding layer in the decoder to obtain a new feature map with effective attention information. In the decoder, the feature map of each layer is passed twice through DSC to reduce the number of channels and an upsampling technique is used to expand its size. Finally, the generated feature map is converted into the output image by 1 × 1 convolution. The discriminator uses the same network architecture as GANomaly (Lecun et al., 1998) to extract the latent features of the input image and distinguish whether the image is real or generated.

3.3 Model training process

Although various modifications have been made, the loss function utilised in this model is somewhat similar to the Skip-GANomaly model. Adversarial loss, context loss, and potential loss are still the three components of the loss function.

The definition of adversarial loss is predicated on the antagonistic interaction between the discriminator and the generator. We want the discriminator to consider the image produced by the generator to be a real image; in other words, we want the discriminator's probability value to be near 1. The discriminator should be able to reliably differentiate between created and genuine images; that is, it should produce a probability value that is near to 0 for generated images and near to 1 for real photos. The expression in mathematics is:

$$L_{adv} = -E_{x \sim Pdata(x)} \left|\log D(x)\right| - E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (5)$$

where the distribution of random noise is denoted by $p_x(z)$ and the distribution of actual data by $p_{data}(x)$.

Context loss: designed to quantify how much context information is lost between images produced by the generator and actual photographs. Contextual information in an image encompasses the general arrangement, specific details, and the connections among various elements. The model's performance will suffer if the generator's image differs significantly from the original image in terms of context information. The expression in mathematics is:

$$L_{lat} = \frac{1}{N} \sum_{i=1}^{N} \left| h_x(i) - h_{G(z)}(i) \right|^2 \quad (6)$$

where N is the number of pixels of the image, x_i and $G(z)_i$ are the i^{th} pixel values of the real and generated images, respectively.

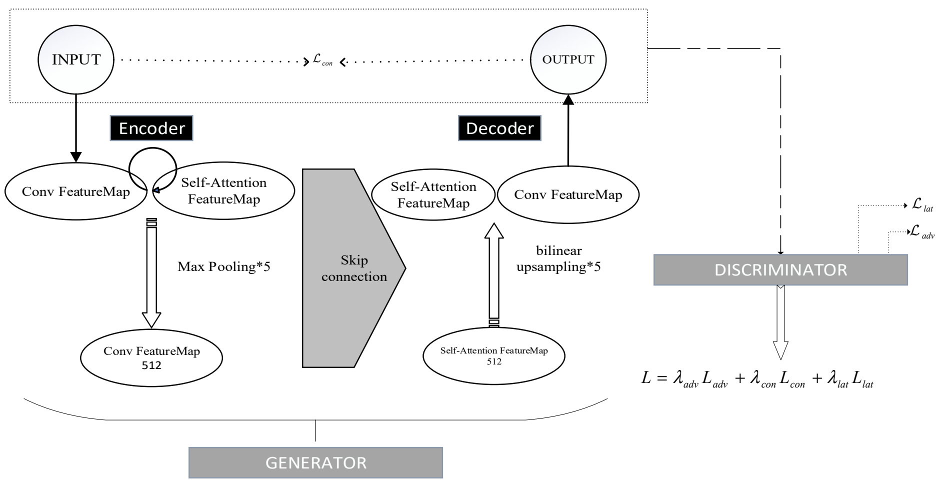
Latent loss: this quantifies how much the generator-generated image differs from the actual image in terms of latent space representation. Because the discriminator and generator in the Skip-GANomaly model both use the latent space, the model's effectiveness depends on how well the latent space is represented. The expression in mathematics is:

$$L_{lat} = \frac{1}{N} \sum_{i=1}^{N} \left| h_x(i) - h_{G(z)}(i) \right|^2 \quad (7)$$

where N is the dimension of the latent space representation, $h_x(i)$ and $h_{G(z)}(i)$ are the i^{th} element of the latent representation of the real and generated images, respectively.

In this study, the receiver operating characteristic (ROC) curve area under the curve (AUC) was chosen as the main evaluation metric, while also considering metrics such as accuracy, recall, and F1-score. AUC is an effective method to evaluate the detection ability of binary detection models, and is also widely used as a model evaluation method for deep learning. The specific method is to preprocess the normal images and abnormal images respectively, and obtain the corresponding output results through the model. Accuracy refers to the proportion of correctly predicted samples to the total number of samples, while recall refers to the proportion of actual positive samples among the samples predicted as positive by the model. The F1-score is the harmonic mean of accuracy and recall, which comprehensively considers both indicators and can more comprehensively evaluate the performance of the model. The specific process is shown in Figure 5.

Figure 5 Specific flowchart



3.4 *Setting training objectives*

The goal of the training process is to optimise the model's performance in the anomaly detection job by modifying the discriminator and generator's parameters. In order to fool the discriminator, the generator's images must be as realistic as feasible, and the discriminator must be able to tell the difference between created and genuine images. This calls for adversarial learning, which alternately optimises the discriminator D and generator G in order to minimise the loss function. At the same time, the accuracy of the latent space representation and the context information should be considered in order to further enhance the model's accuracy for anomaly detection. Consequently, the following is an expression of the training objective.

$$\text{minimise } L = w_{adv} L_{adv} + w_{ctx} L_{ctx} + w_{lat} L_{lat} \tag{8}$$

where w_{adv}, w_{ctx} and w_{lat} are the weight parameters for adversarial loss, context loss, and potential loss, respectively. By tuning these weight parameters, the performance of the model can be optimised according to the specific task and dataset.

4 Experimental setup

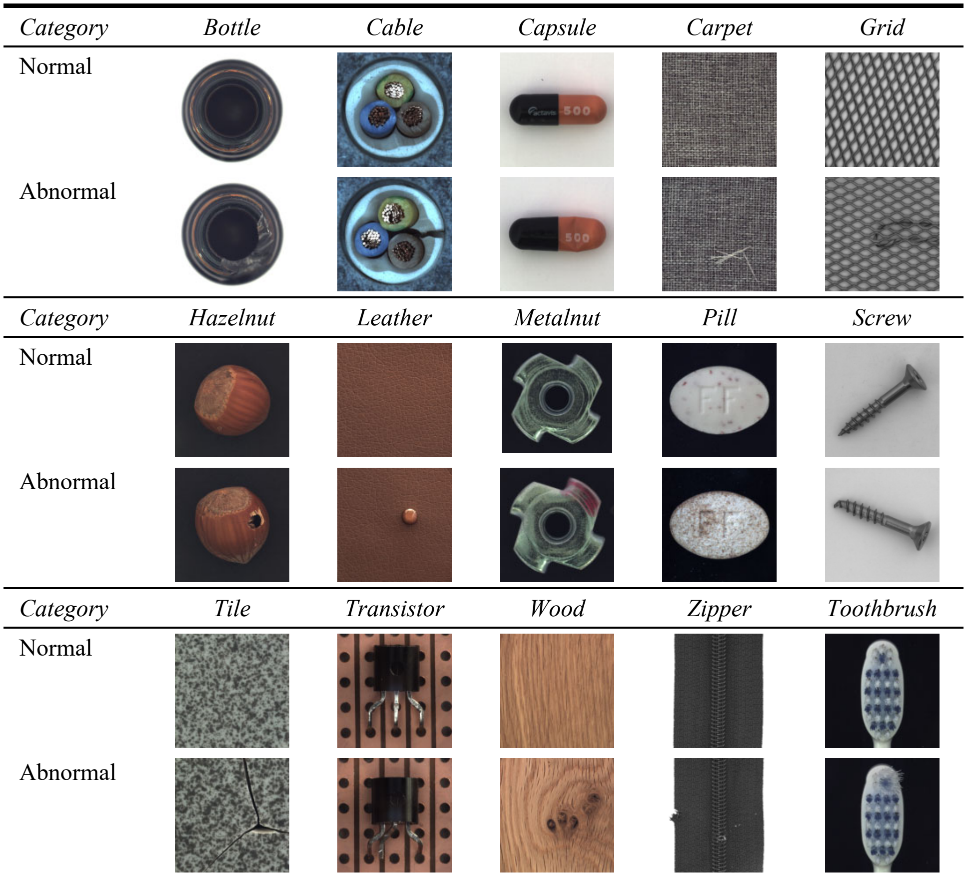
4.1 *Experimental dataset*

4.1.1 *MVTec AD*

The MVTec dataset (Bergmann et al., 2020) is dedicated to visual inspection and contains a total of 5,354 images, which include five textures and ten object categories.

The training set has 3,629 images without defects (normal), the test set has 467 images without defects (normal) and 1,258 images with defects (abnormal), and the resolution of the images is between 700 × 700 and 1,024 × 1,024. The dataset is shown in Figure 6.

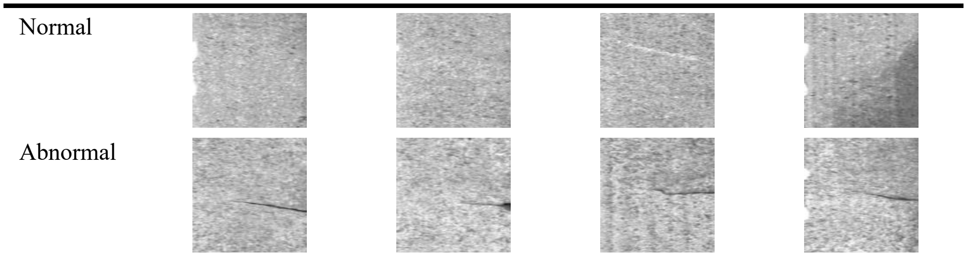
Figure 6 The MVTec AD dataset (see online version for colours)



4.1.2 KolektorSDD

KolektorSDD dataset (Tabernik et al., 2020) integrates the set of defect images of electronic commutators. These images reveal possible minor damage or cracks on the plastic embedding surface on the electronic commutator. These images were captured in a controlled environment with uniform illumination. The defect image contains pixel-level annotations of the defect parts. The dataset contains eight non-overlapping images of each surface of 50 defective electronic commutators, for a total of 399 images including 52 defective and 347 non-defective images as shown in Figure 7.

Figure 7 Partial picture of KolektorSDD dataset



4.2 Experimental platform and network parameter settings

- *The experimental environment is configured as follows:* Intel 13th generation i5-13600KF processor, Nvidia GeForce RTX4070Ti graphics card for GPU operation, 1TB solid state drive and 64GB running memory, Cuda version v12.0.76, cudnn version v8.9.5, and CUDA version v8.9.5. Pytorch is used as the deep learning framework for training and validation. The network runs on the VSCODE platform.
- *Training parameters:* Adam is employed as the optimiser with a learning rate of 0.0002 to guarantee quick and effective training. In equalities, loss functions have been defined to represent prospective loss, context loss, and adversarial loss, respectively. For the adversarial loss (L_{adv}), the weight w_{adv} is set to 0.5, the context loss (L_{ctx}), the weight w_{ctx} is set to 0.3, and the latent loss (L_{lat}), the weight w_{lat} is set to 0.2. Accordingly, the weights of the loss function are set to specific values. The training model's iterations are also set to 10,000 in order to guarantee that the model does not crash during training and to allow for a fair comparison with the chosen model.
- *Evaluation metrics:* in this study, the area under the receiver operating characteristic (ROC) (Ling et al., 2003) curve (AUC) (Akcay et al., 2018; Akay et al., 2019; Ronneberger et al., 2015) was used to evaluate the performance of the detection. AUC is an effective way to measure the ability of a binary detection model, which is also widely used to evaluate deep learning models.
- *Comparison experiment:* this paper will be compared with AnoGAN (Schlegl et al., 2017), GANomaly (Akcay et al., 2018), and Skip-GANomaly (Akay et al., 2019).

5 Experimental results and analysis

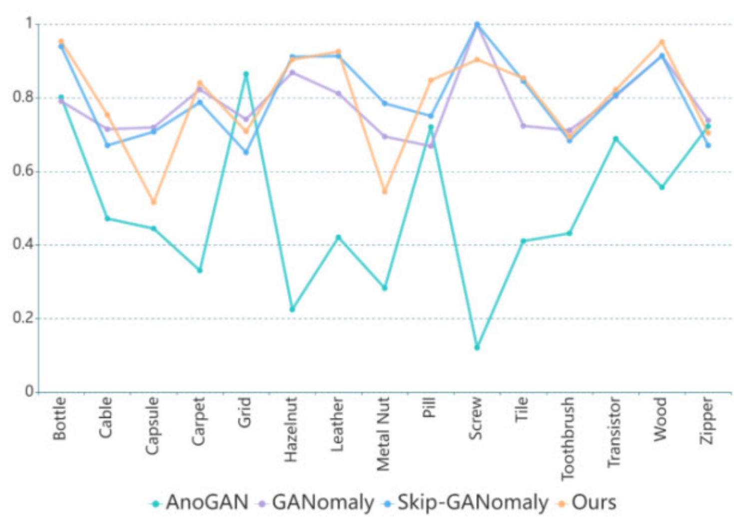
5.1 MVTec AD dataset

As shown in Table 1 and Figure 8, the proposed model has the best performance in 8 categories of 15 categories in MVTec AD. In the other seven categories, except capsule and metal nut have slightly worse performance, the remaining AUCs are not much different from the highest value, or even almost the same.

Table 1 AUC of each category on AnoGAN, GANomaly, Skip-GANomaly, and improved models for MVTec AD dataset

Category	*AnoGAN*	*GANomaly*	*Skip-GANomaly*	*Ours*
Bottle	0.802	0.791	0.940	0.954
Cable	0.472	0.715	0.671	0.754
Capsule	0.445	0.72	0.708	0.516
Carpet	0.331	0.823	0.788	0.841
Grid	0.865	0.742	0.652	0.709
Hazelnut	0.225	0.869	0.912	0.904
Leather	0.421	0.812	0.914	0.926
Metal_Nut	0.283	0.695	0.785	0.545
Pill	0.721	0.669	0.751	0.848
Screw	0.121	1.00	1.00	0.904
Tile	0.411	0.724	0.846	0.854
Toothbrush	0.432	0.712	0.684	0.695
Transistor	0.689	0.805	0.808	0.822
Wood	0.557	0.915	0.914	0.952
Zipper	0.723	0.739	0.671	0.705

Figure 8 Line plot of AUC of MVTec AD dataset on AnoGAN, GANomaly, Skip-GANomaly, and improved model (see online version for colours)



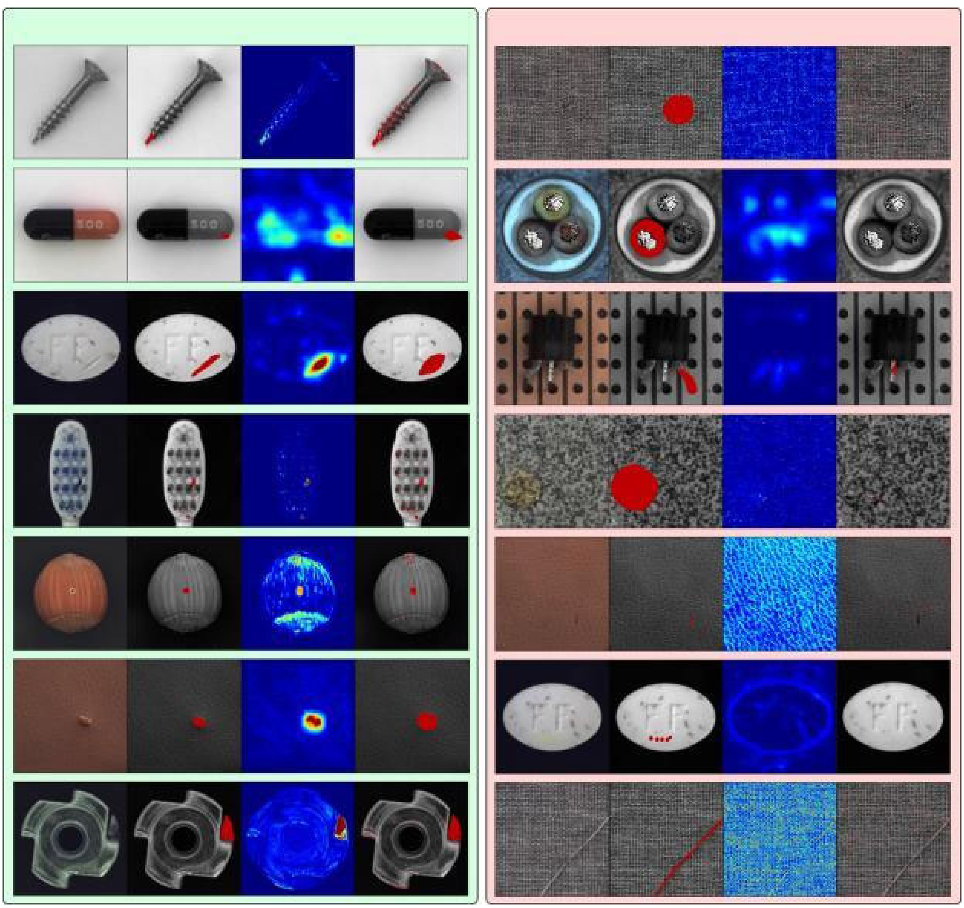
The detection effects in the categories of bottle, carpet, leather, transistor and wood are about 1%–10% higher than those of AnoGAN, GANomaly and Skip-GANomaly, indicating that the performance of the model has been improved. The AUCs of capsule, metal nut and zipper categories are obviously low, and the main reasons are as follows:

1 Data characteristics and model adaptability: although self-attention is able to capture long-distance dependencies, the MVTec AD dataset contains a variety of industrial product images. Due to the complex spatial distribution of features of several classes with poor performance, it is difficult for the model to effectively learn and represent these features.

2 The self-attention module may cause the model to focus too much and ignore other important areas.

3 Intra-class differences are not significant enough: the differences between normal images and abnormal images of capsule, metal nut and zipper are not significant enough, especially metal nut and zipper, which makes it difficult to distinguish even if self-attention is used.

Figure 9 The test results are presented (see online version for colours)



5.2 KolektorSDD

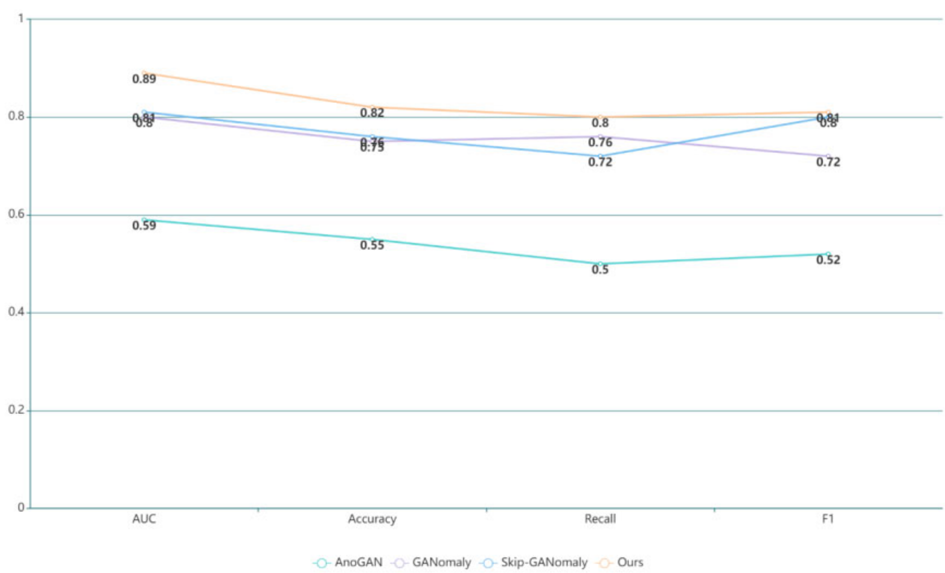
Table 2 lists the results of AUC scores obtained by comparing the proposed method with AnoGAN, GANomaly, and Skip-GANomaly on KolektorSDD dataset.

The results of comparing the suggested approach with AnoGAN, GANomaly, and Skip-GANomaly on the KolektorSDD dataset are shown in Table 2 along with the AUC scores, precision, recall, and F1 value. With a roughly 30% improvement over AnoGAN findings, a 9% improvement over GANomaly, and an 8% improvement over Skip-GANomaly results, the suggested model exhibits a greater detection capacity when dealing with datasets that contain highly complicated changes, such as KolektorSDD.

Table 2 AUC of AnoGAN, GANomaly, Skip-GANomaly, and the improved model on KolektorSDD dataset

Model	*AUC*	*Accuracy*	*Recall*	*F1*
AnoGAN	0.59	0.55	0.50	0.52
GANomaly	0.80	0.75	0.70	0.72
Skip-GANomaly	0.81	0.76	0.72	0.74
Ours	0.89	0.82	0.80	0.81

Figure 10 Experimental result comparison chart of KolektorSDD dataset (see online version for colours)



6 Results and prospects

An enhanced network that is predicated on the Skip-GANomaly model is suggested. Self-attention added to the encoder greatly improves the model's capacity to represent data and the precision of anomaly identification, allowing for a more comprehensive comprehension of the connection between the image's local features and overall structure. Finding and recognising anomalies is made easier by the model's ability to dynamically focus on the most significant areas of the image thanks to the self-attention module. The results also show that, in most categories, the model performs better than the AnoGAN, GANomaly, and Skip-GANomaly network models; however, the model still needs to be improved because its recognition ability for some single background categories is still insufficient, which limits its applicability in real-world settings like factories.

References

Akcay, S., Atapour-Abarghouei, A., Breckon, T.P. (2018) *GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training*, Newcastle University, DOI: 10.1007/978-3-030-20893-6_39.

Akay, S., Atapour-Abarghouei, A. and Breckon, T.P. (2019) *Skip-GANomaly: Skip Connected and Adversarially Trained Encoder-Decoder Anomaly Detection*, IEEE, DOI: 10.1109/IJCNN.2019.8851808.

Bergmann, P., Fauser, M., Sattlegger, D. et al. (2020) 'MVTec AD – a comprehensive real-world dataset for unsupervised anomaly detection', *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, DOI: 10.1109/CVPR.2019.00982.

Chen, X. (2004) 'The world machine tool industry present situation, market demand and development trend', *Journal of Precision Manufacturing and Automation*, Vol. 2004, No. 4, p.2, DOI: CNKI: SUN: SCYM.0.2004-04-001.

Gatys, L.A., Ecker, A.S. and Bethge, M. (2018) *Image Style Transfer Using Convolutional Neural Networks Overview of Method Content: Global Structure*, DOI: 10.1109/cvpr.2016.265.

Goodfellow, I., Pouget-Abadie, J., Mirza, M. et al. (2020) 'Generative adversarial networks', *Communications of the ACM*, Vol. 63, No. 11, pp.139–144.

Hu, L., Wang, C., Sun, Z. et al. (2020) 'Image recognition method based on WGAN', *Journal of Control Engineering*, Vol. 27, No. 12, p.8.

Lecun, Y.A., Bottou, L., Orr, G.B. et al. (1998) 'Efficient backprop', *Lecture Notes in Computer Science*, pp.9–48, DOI: 10.1007/978-3-642-35289-8-3.

Lin, J-Q., Chen, S., He, H. et al. (2021) 'High quality CT image based on particular loss function GAN prediction', *Journal of Control Engineering*, DOI: 10.14107/j.carol carroll nki KZGC 20190312.

Ling, C.X., Huang, J., Zhang, H. (2003) 'AUC: a statistically consistent and more discriminating measure than accuracy', *International Joint Conference on Artificial Intelligence*, Morgan Kaufmann Publishers Inc.

Luo, D., Cai, Y., Yang, Z. et al. (2022) 'Review of deep learning methods for industrial defect detection', *Science China Information Science*, Vol. 52, No. 6, pp.1002–1039.

Liu, G., Lan, S., Zhang, T. et al. (2021) *SAGAN: Skip-Attention GAN For Anomaly Detection*, IEEE, DOI: 10.1109/ICIP42928.2021.9506332.

Radford, A., Metz, L. and Chintala, S. (2015) 'Unsupervised representation learning with deep convolutional generative adversarial networks', *Computer Science*, DOI: 10.48550/arXiv.1511.06434.

Ronneberger, O., Fischer, P. and Brox, T. (2015) *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Springer International Publishing, DOI: 10.1007/978-3-319-24574-4_28.

Schlegl, T., Seebck, P., Waldstein, S.M. et al. (2017) *Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery*, Springer, Cham, DOI: 10.1007/978-3-319-59050-9_12.

Schlegl, T., Seebck, P., Waldstein, S.M. et al. (2019) 'f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks', *Medical Image Analysis*, Vol. 54, DOI: 10.1016/j.media.2019.01.010.

Sun, Y., Yang, Z., Tao, B. et al. (2021) 'Multiscale generative adversarial network for real-world super-resolution', *Concurrency and Computation: Practice and Experience*, DOI: 10.1002/cpe.6430.

Tabernik, D., Šela, S., Skvarč, J. et al. (2020) 'Segmentation-based deep learning approach for surface defect detection', *Journal of Intelligent Manufacturing*, Vol. 31, No. 3, pp.759–776.

Zhang, H., Goodfellow, I., Metaxas, D. et al. (2018) *Self-Attention Generative Adversarial Networks*, DOI: 10.48550/arXiv.1805.08318.

Zhang, T., Liu, Y., Yang, Y. et al. (2020) 'Review of surface defect detection based on machine vision', *Science Technology and Engineering*, Vol. 20, No. 35, pp.14366–14376.

Zenati, H., Romain, M., Foo, C.S. et al. (2018) *Adversarially Learned Anomaly Detection*, IEEE, DOI: 10.1109/ICDM.2018.00088.

Zhou, L., Wang, Z., Sun, D. et al. (2017) 'Modern precision measurement technology and development', *Journal of Instruments and Meters*, Vol. 38, No. 8, p.10, DOI: 10.3969 /j.issn.0254-3087.2017.08.005.