APEM journal

Advances in Production Engineering & Management
Volume 21 | Number 1 | March 2026 | pp 101–116
https://doi.org/10.14743/apem2026.1.564

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Attention-guided unsupervised representation network for anomaly detection in electronic component manufacturing images

Liu, B.[a,*], **Li, C.S.**[b]

[a]Chengdu Aeronautic Polytechnic University, Chengdu, P.R. China
[b]Geely University of China, Chengdu, P.R. China

ABSTRACT

In modern industrial production, electronic component manufacturing imposes increasingly stringent requirements on quality inspection. To address the susceptibility of traditional detection methods to noise interference, the difficulty of localizing abnormal regions in images, and insufficient feature extraction capability, an improved industrial image anomaly detection method based on the original AGUR-Net (Attention-Guided Unsupervised Representation Network) model is proposed. This method enhances the recognition of abnormal regions by introducing attention gate modules and preactivated fusion residual blocks, while simultaneously combining stacked sparse denoising autoenco-ders to enhance the model's robustness to noise and suppress redundant information. The experimental results show that the proposed method achieves an accuracy of 84.26 % in locating abnormal regions on the MVTecAD dataset after 185 iterations, which is higher than the dual attention generative adversarial network (79.21 %), long short-term memory non-destructive testing network (79.31 %), and bidirectional long short-term memory network (81.63 %). At 400 iterations, the average loss value of this method for detecting abnormal images of industrial products is 0.016. In addition, the processing time for single images of Class A, B, and C defects in the actual factory environment is 372 ms, 329 ms, and 378 ms, respectively, demonstrating high detection efficiency. Overall, this method can accurately identify and locate abnormal images in the surface treatment and quality inspection stages of electronic component production, which helps to improve the quality control level in the manufacturing process and provides effective technical support for intelligent industrial production.

ARTICLE INFO

Keywords:
Electronic component manufacturing;
Industrial anomaly detection;
Unsupervised anomaly detection;
Anomaly localization;
Attention gate;
Stacked sparse denoising autoencoder;
Multimodal generative adversarial network (MGAN)

*Corresponding author:
liubing@cap.edu.cn
(Liu, B.)

Article history:
Received 21 July 2025
Revised 25 April 2026
Accepted 26 April 2026

1. Introduction

Image anomaly detection is a computer vision technology that meets practical application needs in industrial quality control, medical diagnosis support, and remote sensing with geographic information systems [1]. In industrial production, it helps detect appearance defects on products, such as tiny scratches on semiconductor chips, bubbles in glassware, and dents on screws [2, 3]. By using image anomaly detection for real-time quality monitoring, production efficiency can be improved. Traditional image detection methods for industrial products include threshold-based and edge-based techniques [4]. These methods can quickly detect shape changes and cracks in images. However, they are sensitive to noise and often miss surface defects on products with rich textures [5]. To address the difficulty of accurate defect identification, Ullah *et al*.

proposed an intelligent defect chip detection framework and introduced an inverse feature matching mechanism to improve the interpretability of abnormal region localization. Their experiments showed that the method effectively distinguished fine-grained defect details in chips [6]. To solve the time consumption problem in bearing defect detection, Liu *et al*. proposed an integrated unsupervised learning method based on autoencoder and generative adversarial networks. The results showed high detection efficiency, making the method suitable for real-time applications [7]. Image anomaly detection also plays an important role in part segmentation during industrial production. Threshold segmentation divides image pixels into categories and separates the target from the background to locate abnormal regions effectively [8]. To address challenges in detecting complex anomalies, Bergmann *et al*. proposed a logic-constrained method for unsupervised anomaly detection and localization. Experimental results confirmed that the method reduced ambiguity in locating logical anomalies [9]. However, many of these detection techniques are still highly affected by noise and environmental interference, making the detection results unstable. A new method is needed to improve feature recognition in industrial product images.

As the scale of industrial production continues to grow, image anomaly detection is expected to offer higher detection accuracy and more precise localization [10]. In response, researchers have explored various algorithms. Zhang *et al*. proposed a deep unsupervised learning method based on an improved adversarial autoencoder to solve the problem of data imbalance in anomaly detection. Their experiments showed that the method detects diseased samples by identifying significant deviations from the normal distribution under unsupervised conditions [11]. Sun *et al*. introduced a stochastic anomaly multi-scale feature focusing autoencoder network to reduce the missed detection of abnormal samples. The method successfully solved the problem of blurry image reconstruction [12]. Although these image anomaly detection algorithms can effectively extract abnormal features, they require long inference times and substantial computing resources, and thus cannot be readily applied to the detection of subtle abnormal areas in industrial commodities from raw material preparation to packaging and transportation [13]. Therefore, in view of the difficulty in detecting local abnormal image areas at each stage of electronic component production management, this study proposes an abnormal image detection method for electronic component production based on the Attention-Guided Unsupervised Representation Network (AGUR-Net). This method improves the feature extraction ability for abnormal areas in industrial commodity images. This method can precisely locate abnormal local areas such as short circuits and open circuits on the circuit board, ensuring that the raw materials put into production are qualified in the quality inspection of raw materials for electronic component production. Not only that, using the AGUR-Net industrial image anomaly detection algorithm to collect images of randomly selected electronic components on the production line and analyze the internal structure can further ensure that the product quality meets the standards. The concept of "unsupervised" needs to be clearly defined: this method only uses normal sample images during the training phase and does not rely on any position or category labeling of abnormal samples, thus belonging to unsupervised learning in terms of method properties. However, in the performance evaluation stage of the model, in order to quantitatively measure the accuracy of anomaly localization and classification, the experiment used a standard dataset with category labels and category-based evaluation indicators. This "unsupervised training, supervised evaluation" model is quite common in the field of industrial anomaly detection and does not change the unsupervised nature of the model training process. It is expected that this method can reduce the occurrence of false inspections in industrial quality inspection and provide a more comprehensive and intelligent solution for industrial production. Unlike the original AGUR-Net model, the main contribution of this paper lies in two aspects: first, an improved attention gate module and a pre-activated fusion residual block are introduced in the U-shaped decoding path, replacing the traditional skip connection method and allowing the network to dynamically focus on potential abnormal areas rather than uniformly processing all spatial positions [14]. Second, the stacked sparse denoising autoencoder and multimodal generative adversarial learning module are embedded into the front-end feature processing flow of AGUR-Net to suppress noise interference during image acquisition and enhance multimodal feature expression. The above com-

ponents are all designed for abnormal detection scenarios in electronic component production in this study, rather than being directly transferred from existing literature. In contrast, the hollow space pyramid pooling module and efficient channel attention module used in the model are based on existing research, and their main function is to provide multi-scale contextual information and channel weight calibration for the network. By combining the new and old components mentioned above, the proposed method can improve the localization accuracy and detection robustness of abnormal areas while maintaining computational efficiency.

2. Methods and materials

2.1 Design of the anomaly image detection algorithm based on AGUR-Net

With continued technological progress, image anomaly detection is being applied increasingly widely in industry, security, and healthcare [15, 16]. This technology can effectively detect product defects such as scratches and cracks, and also identify abnormal behaviors and suspicious objects. Although traditional detection methods can detect abnormalities, they are easily affected by noise and local changes, and their performance is unsatisfactory [17, 18]. To address the difficulties of anomaly localization and insufficient feature extraction, an AGUR-Net-based anomaly image detection algorithm is proposed to improve detection efficiency. The overall framework of the algorithm is shown in Fig. 1.

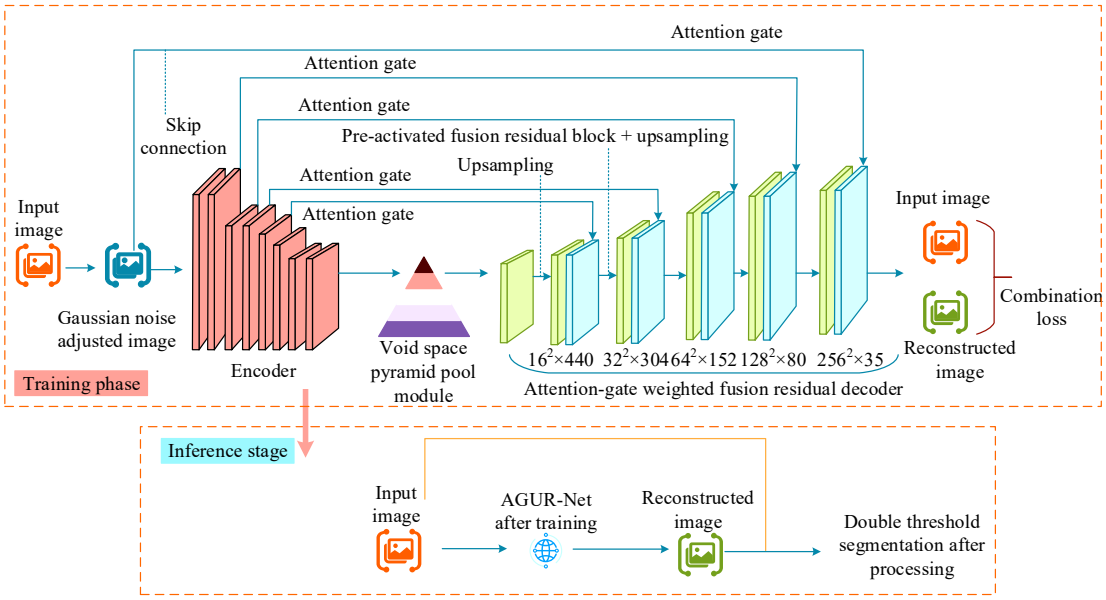


Fig. 1 Image anomaly detection algorithm framework based on AGUR-Net

As shown in Fig. 1, the framework of the AGUR-Net-based anomaly detection algorithm includes a training phase and an inference phase. In the training phase, the input image is first sent into the network and adjusted by introducing Gaussian noise. Then, the adjusted image passes through the encoder to extract features while gradually reducing the size of the feature maps. Next, the Atrous Spatial Pyramid Pooling (ASPP) module processes the encoded features to obtain multi-scale context information. During this process, attention gates in the residual decoder are used to enhance key features through weighted fusion. The encoder and attention gates reconstruct the structure and details of the image to optimize training. In the inference phase, the test image is input into the trained AGUR-Net. A double-threshold segmentation post-processing method is applied to the reconstruction result to locate and identify anomaly regions. The calculation of the atrous convolution in the improved ASPP module is shown in Eq. 1.

$$f'[i] = \sum_{k} f\,[i + r \cdot k]W[k] \tag{1}$$

In Eq. 1, i and f represent the position and input feature, $f'[i]$ and $W[k]$ represent the output feature map and kernel convolution, and r denotes the sampling stride of the atrous convolution. The attention gate coefficient is calculated as shown in Eq. 2 [19].

$$\alpha_i^l = \sigma_2\left(q_{att}^l(x_i^l, g_i; \Theta_{att})\right) \tag{2}$$

In Eq. 2, q_{att}^l is the intermediate variable in the attention module, g_i is the gating signal, x_i^l is the input feature vector of layer l, σ_2 is the Sigmoid activation function, and Θ_{att} is the set of all learnable parameters in the attention module. The improved attention gate module and pre-activation residual block are shown in Fig. 2.

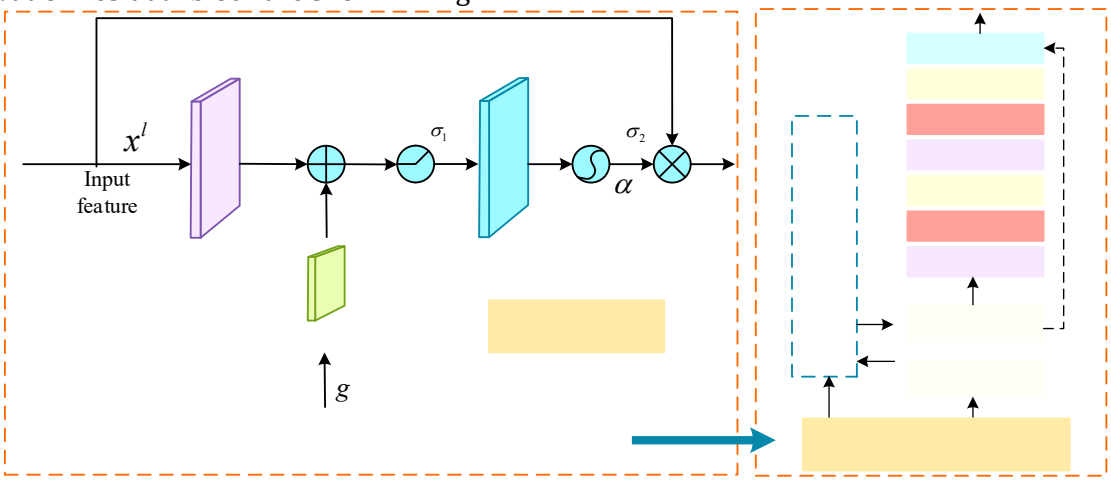


Fig. 2 Improved attention gate module and pre-activation fusion residual block

As shown in Fig. 2, the attention gate module receives input feature x^l and gate feature g. These features are separately processed through 1×1 convolution layers to adjust dimensions and are then fused. The fused feature passes through a ReLU activation function to enhance feature representation. Finally, the Sigmoid function generates the attention coefficient, which is multiplied with x^l to highlight important features. In the pre-activation residual block, channel-wise concatenation is done through upsampling. Then, the attention gate mechanism adaptively adjusts feature weights. After that, the convolution output is normalized and passed through the ReLU function to introduce non-linearity and complete the residual connection. To perform deeper anomaly detection, the algorithm introduces a deep anomaly scoring network based on AGUR-Net to further distinguish abnormal image data. This network learns complex features automatically and adapts well to high-dimensional data [20, 21]. The anomaly score is calculated as shown in Eq. 3.

$$\xi(f_{ij}, \theta_s) = \sum_{n=1}^{L} w_n^o f_{ijk} + w_{L+1}^o \tag{3}$$

In Eq. 3, f_{ij} and θ_s are the feature vector and weight, w_{L+1}^o is the anomaly score bias term, and k is the number of terms. The anomaly score mapping is calculated as shown in Equation (4).

$$\psi(x_{ij}, \theta) = \xi(\varphi(x_{ij}; \theta_f); \theta_s) \tag{4}$$

In Eq. 4, ψ and θ_f represent the anomaly score mapping and weight matrix. The global compression feature in the improved Efficient Channel Attention (ECA) module is calculated as shown in Eq. 5.

$$y = \frac{1}{W \times H} \sum_{v=1}^{W} \sum_{b=1}^{H} x_{(e,r)} \tag{5}$$

In Eq. 5, H is the height, W is the width of the feature map, $x_{(v,b)}$ is the feature value, and y is the global compression feature. The improved ECA module first inputs a feature map with height

H, width W, and channel number C. Then global pooling is applied to compress the spatial dimensions, producing an output of size $1 \times 1 \times C$. The globally pooled channel information is mapped to capture dependencies between channels. The result is passed through the Sigmoid function σ to obtain attention weights with output dimension $1 \times 1 \times C$. These weights are multiplied with the corresponding channels of the original feature map for channel-wise weighting. The final output is calculated as shown in Eq. 6.

$$\hat{x} = W_z \cdot x \tag{6}$$

In Eq. 6, W_z is the important channel weight. W_z is calculated as shown in Eq. 7 [22].

$$W_z = \sigma\big(D_r(y)\big) \tag{7}$$

In Eq. 7, r is the kernel size for one-dimensional convolution, and D_r is the one-dimensional convolution result. By calculating the importance of each channel, the overall detection efficiency improves while ensuring parameter sharing. Channel information is calculated as shown in Eq. 8.

$$z_c^1 = F_{GAP}(F_c) = \frac{1}{H \times W} \sum_{v=1}^{H} \sum_{b=1} F_c\,(v, b) \tag{8}$$

In Eq. 8, z_c^1 is the channel information, and F_c represents the feature map after channel fusion of c. In conclusion, the deep anomaly scoring network and AGUR-Net can accurately capture anomaly data and detect industrial product defects. The improved AGUR-Net structure is shown in Fig. 3.

It can be seen from Fig. 3 that the improved AGUR-Net anomaly image detection algorithm consists of three modules: front-end feature extraction and processing, deep anomaly scoring network, and back-end reconstruction and post-processing. The front-end module encodes using a U-shaped network, receives input features through the attention gate module, combines 1×1 convolution and feature fusion to enhance the feature learning ability, and refines the processing through multi-scale feature fusion. The deep anomaly scoring network maps the processed features to the anomaly scoring space and calibrates the anomaly scores to determine the location and degree of anomalies. Finally, the back-end module restores the image features through the decoder and locates and segments the abnormal areas by using operations such as double-threshold segmentation. This algorithm can detect abnormalities in real time during processes such as the assembly and soldering of electronic components, and promptly issue alerts to feed back to the production line control system.

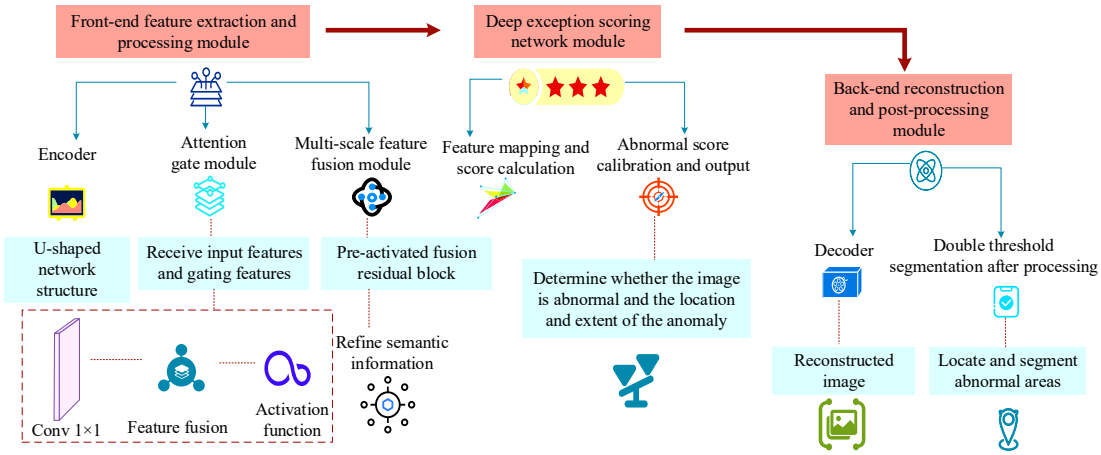


Fig. 3 The improved AGUR-Net abnormal image detection algorithm structure

2.2 Optimization of the detection method based on AGUR-Net and SSDAE

To address the limitations of the AGUR-Net algorithm in detecting minor defects in electronic components, as well as the vulnerability of traditional methods to noise interference, this study proposes an improved method integrating SSDAE. By enhancing the anti-noise ability and extracting key features, the detection accuracy of micro-defects in industrial scenarios is improved. SSDAE suppresses redundant information and reduces noise during image collection [23]. The improved SSDAE network structure is shown in Fig. 4.

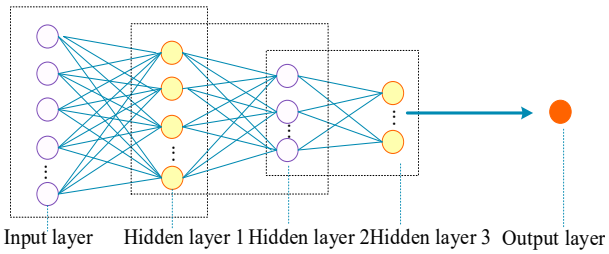


Fig. 4 Improved SSDAE network structure

As shown in Fig. 4, the SSDAE structure consists of input, hidden, and output layers. The input layer includes multiple purple circles, each representing an input neuron. It receives raw data and sends it to Hidden Layer 1. Hidden Layer 1 neurons are trained with the input neurons to learn patterns and features. Then, the output of Hidden Layer 1 is used as the input of Hidden Layer 2 for further processing. Hidden Layer 3 performs deeper extraction on the features from Hidden Layer 2 to enhance the feature representation for industrial anomaly detection. The final output layer returns denoised images after training. The compression process is shown in Eq. 9.

$$K_q^l = f_{sigmoid}\left(z_q^l\right) = f\left(\sum_{a=1}^{n} W_{ql}^{l-1} \cdot x_a^{l-1} + b_q^{l-1}\right) \tag{9}$$

In Eq. 9, l and q are the current hidden layer index and number of neurons, a is the number of neurons in the previous layer, and W is the feature weight coefficient. The sparse autoencoder loss function is calculated as shown in Eq. 10.

$$K_L\left(\rho \| \hat{\rho}_q \|\right) = p\log\frac{\rho}{\hat{\rho}_q} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_q} \tag{10}$$

In Eq. 10, $\hat{\rho}_q$ is the average sparse activation value, $K_L\left(\rho \| \hat{\rho}_q \|\right)$ is the relative entropy distance, and ρ is the sparsity parameter. $\hat{\rho}_q$ is calculated as shown in Eq. 11.

$$\hat{\rho}_q = \frac{1}{m}\sum_{a}^{m}\left[n_q(x_a)\right] \tag{11}$$

In Eq. 11, n_q is the activation state of the hidden neurons, and x_a is the unlabeled actual anomaly data. In addition, this study applies the Multimodal Generative Adversarial Learning Algorithm (MGAN) to improve multimodal training effects such as color and texture. MGAN enhances feature expression and reduces missed and false detections [24]. The MGAN structure is shown in Fig. 5.

As shown in Fig. 5, MGAN takes spatial-domain and frequency-domain samples as input. The spatial-domain sample is raw data, and the frequency-domain sample is processed through Fourier transform. These samples are sent to corresponding generators for encoding and feature extraction. Then, the decoders generate images similar to the original spatial and frequency distributions. The discriminator judges whether the input is an anomaly or a near-normal sample generated by the generator. The reconstruction error of the spatial branch is calculated as shown in Eq. 12.

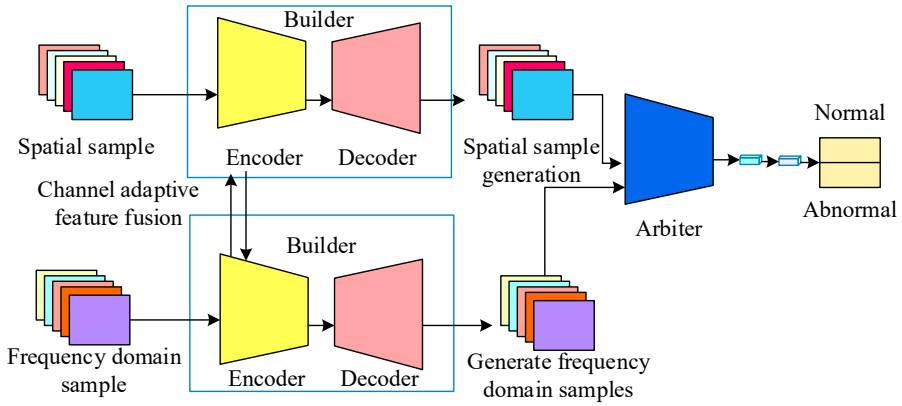


Fig. 5 Schematic diagram of MGAN structure

$$L_{conS} = E_{I\sim p_n}\left[I - \hat{I}_S\right]_1 \tag{12}$$

In Eq. 12, S is the spatial-domain sample and I is the spatial image. The reconstruction error of the frequency branch is calculated as shown in Eq. 13 [25].

$$L_{conF} = E_{I\sim p_n}\left[I - \hat{I}_F\right]_1 \tag{13}$$

In Eq. 13, F is the frequency-domain sample. The adversarial loss is calculated as shown in Eq. 14.

$$L_{adv} = \mathbb{E}_{\hat{I}_{12}\sim p_r}\left[D\left(\hat{I}_{12}\right)\right] - \mathbb{E}_{\hat{I}_3\sim p_r}\left[D\left(\hat{I}_{13}\right)\right] \tag{14}$$

In Eq. 14, $\hat{I}_{12}$ and $\hat{I}_{13}$ are the pixel values of the frequency and spatial branches, and p_r and $D(\cdot)$ are the probability distribution and discriminator function. The improved Fourier transform is calculated as shown in Eq. 15.

$$U(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cdot e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{15}$$

In Eq. 15, $U(u,v)$ is the pixel value in the frequency domain, and u and v are the frequency coordinates. In conclusion, MGAN and SSDAE detect fine features such as texture. The framework based on AGUR-Net for industrial anomaly detection is shown in Fig. 6.

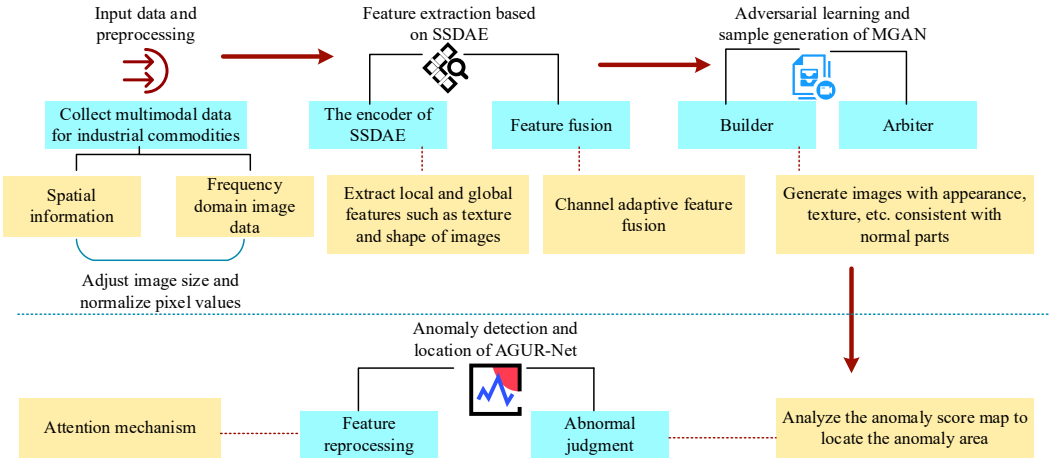


Fig. 6 Workflow of AGUR-Net industrial commodity abnormal image detection method

As can be seen from Fig. 6, the abnormal image detection method for electronic component production based on AGUR-Net first collects and standardizes multimodal data, such as spatial domain and frequency domain image data. Then, the local and global features of the image are extracted using the SSDAE feature extraction module, and adaptive feature fusion is carried out. Then, images similar to normal parts are generated through MGAN adversarial learning to de-

termine the authenticity of the images. Finally, AGUR-Net is used to further detect and locate the abnormal features and output the abnormal score image. This method can regularly inspect the components in the warehouse management of electronic parts and promptly detect surface abnormalities such as discoloration and spots.

3. Results

3.1 Validation of the effectiveness of the AGUR-Net anomaly image detection algorithm

To verify the performance advantages of the AGUR-Net anomaly image detection algorithm, it was compared with three detection algorithms: Dual-Attention Generative Adversarial Network (DA-GAN), Long Short-Term Memory-Non-destructive Testing (LSTM-NDT), and Bi-directional Long Short-Term Memory (BiLSTM). Among them, LSTM-NDT and BiLSTM have been widely applied in industrial image anomaly detection tasks in recent years. The core idea is to use image pixel sequences or feature map sequences as temporal inputs, and utilize the memory ability of recurrent neural networks to capture long-range dependencies and anomaly patterns in images. Therefore, these two methods can serve as reasonable comparison baselines for image-based anomaly detection tasks. DAGAN is an image anomaly detection method based on generative adversarial architecture, which is closer to the task setting of our research method. By comparing multiple methods simultaneously, the performance of AGUR-Net in industrial image anomaly detection can be more comprehensively evaluated. The experiments were conducted on Ubuntu 20.04 with an NVIDIA Tesla V100 GPU (16 GB memory), and all code was implemented using the PyTorch framework. The experiments used the MVTecAD and KSDD2 datasets for training and testing. The MVTecAD dataset contained 5,354 high-resolution color images that covered various industrial anomalies such as scratches and dents. The KSDD2 dataset included 3,336 high-resolution images with 356 labeled defect images. The morphology and texture of abnormal samples in the MVTecAD dataset are highly similar to typical defects on the surface of electronic components, and are therefore used to evaluate the model's ability to locate and detect anomalies related to electronic manufacturing. The data in the KSDD2 dataset is directly related to the surface defect detection tasks of metal and plastic components in electronic component production, and is therefore used to validate the performance of the model on actual electronic manufacturing parts. All test images were resized to 256×256, and the batch size was set to 16. Four anomaly detection algorithms were tested for anomaly region localization, and the results are shown in Fig. 7.

As shown in Fig. 7(a), on the MVTecAD dataset, AGUR-Net achieved an anomaly region localization accuracy of 84.26 % after 185 iterations. At the same iteration count, DAGAN reached 78.63 %. According to Fig. 7(b), on the KSDD2 dataset, AGUR-Net reached a stable accuracy range of 82 % to 96 % after 150 iterations. At 150 iterations, LSTM-NDT and BiLSTM achieved localization accuracies of 77.26 % and 81.96 %, respectively. These results indicated that AGUR-Net effectively located defect regions in industrial quality inspection tasks and demonstrated good applicability. To further verify the accuracy of AGUR-Net in anomaly image detection, it was compared with DAGAN, LSTM-NDT, and BiLSTM on industrial anomaly detection accuracy. The results are shown in Fig. 8.

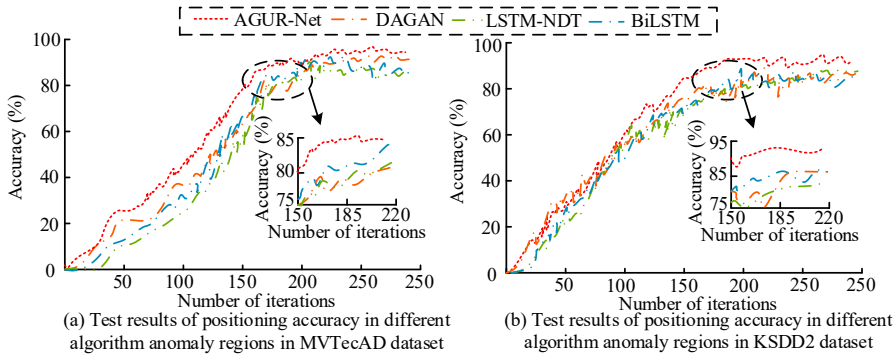


Fig. 7 Anomaly region localization accuracy test results

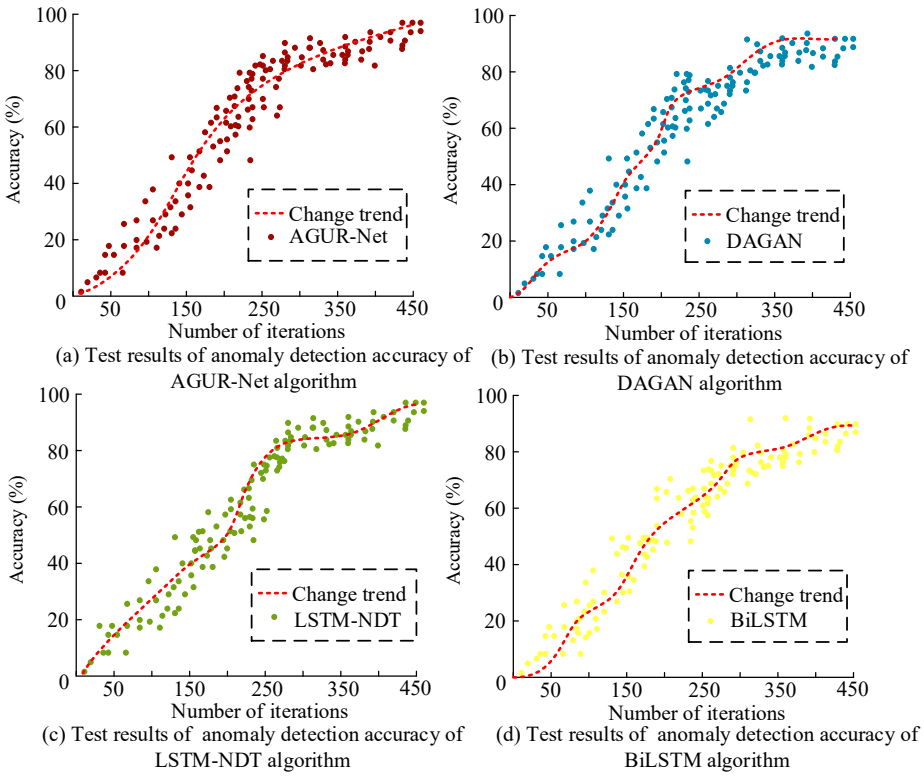


Fig. 8 Anomaly detection accuracy test results

Fig. 8(a) shows that AGUR-Net achieved an industrial anomaly detection accuracy of 43.69 % at 150 iterations. After 250 iterations, the accuracy exceeded 75 %, peaking at 97.26 %. Fig. 8(b) shows that the accuracy of DAGAN fluctuated significantly between 100 and 250 iterations, reaching 73.12 % at 300 iterations. Fig. 8(c) shows that LSTM-NDT maintained an accuracy range of 75 % to 82 % between 250 and 350 iterations. Fig. 8(d) shows that BiLSTM achieved a maximum accuracy of 76.31 % at 275 iterations. Overall, AGUR-Net demonstrated higher detection accuracy and ensured safety in the manufacturing of critical components such as those in aerospace. To further evaluate the classification performance, the algorithm was tested alongside DAGAN, LSTM-NDT, and BiLSTM for anomaly classification, as shown in Fig. 9.

Fig. 9(a) shows that without classification, capsule, screw, leather, and canvas anomaly images were scattered. Fig. 9(b) shows that after classification by AGUR-Net, canvas anomaly images clustered around the coordinate (-3, 3). However, a few canvas images were mistakenly classified into the capsule region. Fig. 9(c) shows that DAGAN struggled with capsule classification, with some being misclassified as leather. Fig. 9(d) shows that LSTM-NDT achieved good classification results for screw and leather anomalies, clustering around (2, -2) and (4, 2), respectively. Fig. 9(e) shows that BiLSTM performed well in classifying leather anomalies, with most clustering around (2, 3). These results indicated that AGUR-Net effectively classified different types of anomalies, helping users quickly identify defect types. Furthermore, AGUR-Net, DAGAN, LSTM-NDT, and BiLSTM were evaluated using the Area Under the Curve (AUC) metric. The anomaly rate in the datasets was set to 10 % to ensure sufficient anomaly samples. To enhance result diversity, the Fashion-MNIST dataset was added. It contained images from 10 categories of fashion items such as coats and shirts with diverse textures and shapes, making it suitable for classification tasks. The Fashion-MNIST dataset, although not directly derived from industrial electronic component production scenarios, has diverse texture, shape, and edge features. The purpose of introducing this dataset is not to simulate the actual detection task of electronic components, but to evaluate the generalization ability and feature discrimination ability of various algorithms on a wider range of image distributions. The results are shown in Table 1.

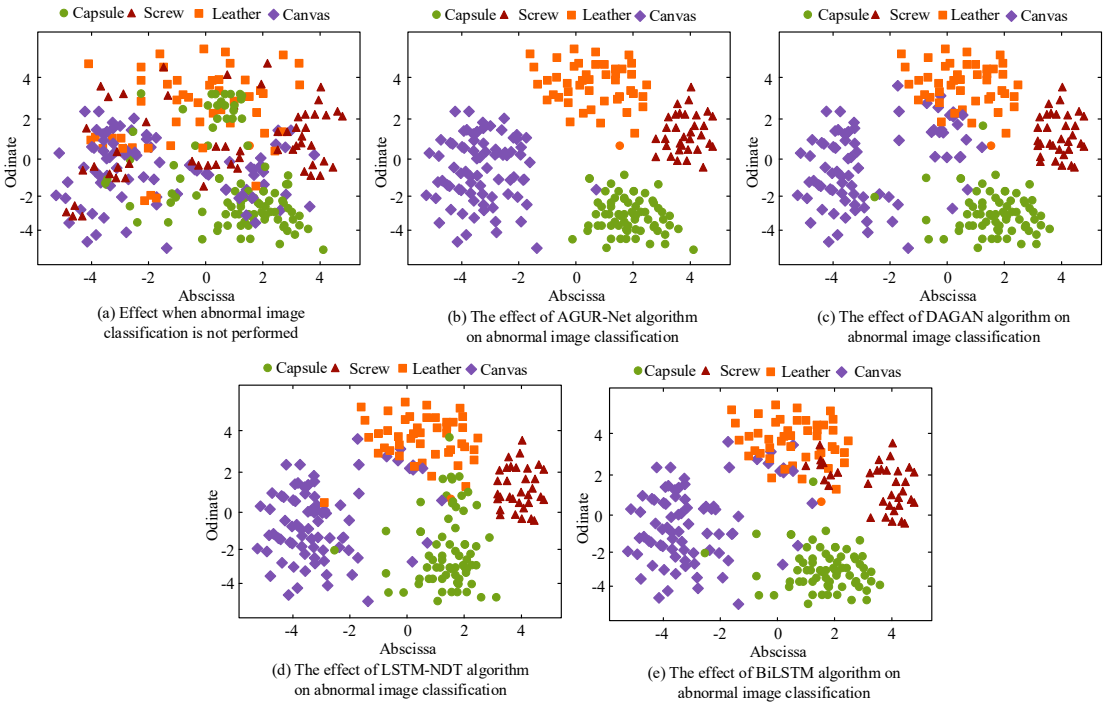


Fig. 9 Anomaly classification results of different detection algorithms

Table 1 Anomaly detection AUC test results of different data sets

Dataset	Correct class	AGUR-Net	DAGAN	LSTM-NDT	BiLSTM
MVTecAD	Bottle	0.836	0.812	0.769	0.746
	Cable	0.923	0.894	0.816	0.836
	Hazelnut	0.946	0.832	0.874	0.816
	Transistor	0.874	0.798	0.816	0.803
KSDD2	Gear wheel	0.913	0.826	0.813	0.781
	lock	0.936	0.864	0.885	0.752
	Plastic case	0.904	0.829	0.836	0.811
	Chip	0.876	0.864	0.814	0.809
Fashion-MNIST	Bag	0.896	0.821	0.838	0.823
	Sandal	0.937	0.903	0.726	0.874
	Pullover	0.948	0.897	0.798	0.893
	Coat	0.968	0.814	0.823	0.846

As shown in Table 1, all four algorithms achieved AUC values above 0.750 on the MVTecAD dataset. For cable samples, AGUR-Net achieved an AUC of 0.923, which was 0.029 higher than DAGAN's 0.894. On the KSDD2 dataset, AGUR-Net achieved an AUC of 0.936 for lock samples, outperforming DAGAN (0.864), LSTM-NDT (0.885), and BiLSTM (0.752). On the Fashion-MNIST dataset, AGUR-Net, DAGAN, LSTM-NDT, and BiLSTM achieved AUC values of 0.937, 0.903, 0.726, and 0.874, respectively, for sandal samples. These results demonstrated that AGUR-Net had a strong ability to distinguish anomalies and maintained high generalization.

3.2 Validation of the AGUR-Net industrial product anomaly image detection method

After verifying the performance of the AGUR-Net anomaly image detection algorithm, the proposed method was further compared with three anomaly image detection methods for electronic component production, namely DAGAN, LSTM-NDT, and BiLSTM. The experiments were conducted on PyTorch 1.8.1, with an initial learning rate of 0.4, and the network was optimized using the Adam optimizer. To ensure the authenticity and reliability of the experiments, the YDFID-1 and ECD datasets were used. The YDFID-1 dataset contained industrial product data such as metal sheets and plastic parts, with 15,000 labeled defects and 3,500 training samples. The data source of the YDFID-1 dataset is relatively close to the raw material and intermediate prod-

uct detection stages in electronic component production, and is used to test the generalization performance of the model on different industrial materials and defect types. The ECD dataset is specifically designed for common electronic components, including high-definition images, and is suitable for object detection, classification, and defect recognition tasks. It is the most suitable test set for practical application scenarios. The visualization results of anomaly location for four abnormal image detection methods of electronic component production are shown in Fig. 10.

Fig. 10(a) shows five abnormal images corresponding to stage 1 (raw material preparation), stage 2 (processing and manufacturing), stage 3 (surface treatment), stage 4 (quality inspection), and stage 5 (packaging and transportation). Fig. 10(b) indicates that the actual abnormal defect is located in the central area of the image in the surface treatment stage. Fig. 10(c) shows that the AGUR-Net-based anomaly image detection method for electronic component production performs better in localizing anomalies in images from the surface treatment stage, and the heat-map distribution is relatively concentrated. Fig. 10(e) shows that the LSTM-NDT abnormal image detection method for electronic component production is not very accurate in the abnormal location of images in the quality inspection stage, and the defect area below the image is not detected. Fig. 10(f) shows that the BiLSTM electronic component production anomaly image detection method has a less effective anomaly location detection effect on images in the processing and manufacturing stage, quality inspection stage, and packaging and transportation stage than the method proposed in the study. Some undamaged areas are detected as abnormal areas. In addition, the study also conducted tests on the average loss values of abnormal electronic component production image detection using four abnormal image detection methods for electronic component production, namely AGUR-Net, DAGAN, LSTM-NDT, and BiLSTM. The test results are shown in Fig. 11.

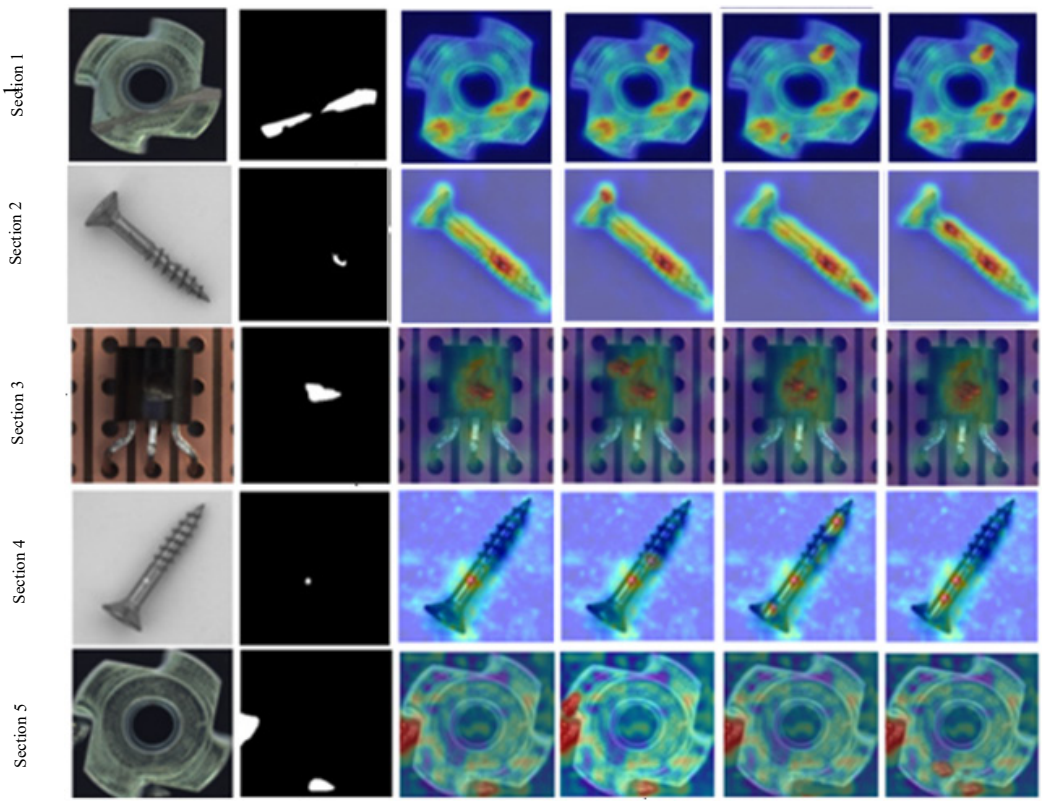
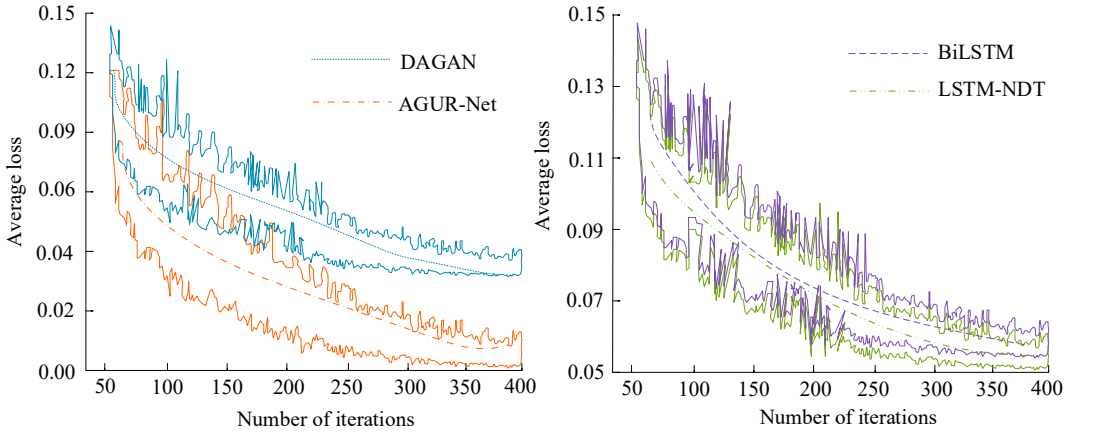


Fig. 10 Comparison of visualization effects of anomaly localization

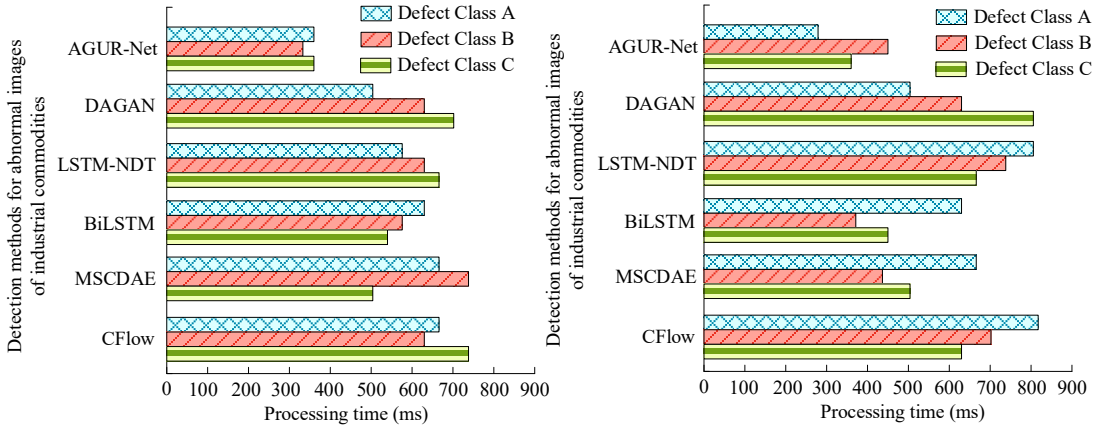


(a) The loss rate of abnormal image detection of industrial commodities by AGUR-Net and DAGAN method

(b) The loss rate of abnormal image detection of industrial commodities by LSTM-NDT and BiLSTM method

Fig. 11 Average loss value test results

Fig. 11(a) shows that AGUR-Net achieved an average loss of 0.016 at 400 iterations. DAGAN had an average loss ranging from 0.05 to 0.07 between 350 and 400 iterations. Fig. 11(b) shows that LSTM-NDT's average loss ranged from 0.053 to 0.082 between 200 and 250 iterations. BiLSTM's loss stabilized after 250 iterations, with a minimum of 0.053. To sum up, the AGUR-Net abnormal image detection method for electronic component production has good training convergence performance and can better capture abnormal features. To further validate its performance, AGUR-Net, DAGAN, LSTM-NDT, and BiLSTM were tested in real-time at Factory X1 and Factory X2. Two additional methods, Multi-Scale Convolutional Denoising Autoencoder (MSCDAE) and Conditional Normalizing Flow (CFlow), were included to increase diversity. These two methods were suitable for tiny defect detection and high-precision anomaly detection. Results are shown in Fig. 12.



(a) Processing time results of single anomaly images with different detection methods in X1 factory

(b) Processing time results of single anomaly images with different detection methods in X2 factory

Fig. 12 Time results of processing a single abnormal image in different data sets

Fig. 12(a) shows that at Factory X1, AGUR-Net processed a single image with type-A, type-B, and type-C defects in 372 ms, 329 ms, and 378 ms, respectively. Fig. 12(b) shows that at Factory X2, the processing times for a single type-A defect image using AGUR-Net, DAGAN, LSTM-NDT, BiLSTM, MSCDAE, and CFlow were 289 ms, 507 ms, 806 ms, 623 ms, 672 ms, and 812 ms, respectively. AGUR-Net achieved faster processing due to its improved attention gate module, which precisely focused on target regions. Overall, AGUR-Net demonstrated high efficiency in real applications. In addition, six methods for detecting abnormal images of electronic component production, namely AGUR-Net, DAGAN, LSTM-NDT, BiLSTM, MSCDAE and CFlow, were also used in the study to test the F1 Score (F1) of abnormal images of industrial electronic component production. The test results are shown in Table 2.

Table 2 F1-score test results in different datasets

Dataset	Method	The number of abnormal samples				
		50	100	150	200	250
YDFID-1	AGUR-Net	0.971	0.939	0.947	0.927	0.935
	DAGAN	0.903	0.893	0.841	0.867	0.836
	LSTM-NDT	0.872	0.810	0.726	0.734	0.712
	BiLSTM	0.813	0.798	0.836	0.874	0.788
	MSCDAE	0.857	0.772	0.802	0.817	0.836
	CFlow	0.826	0.813	0.831	0.846	0.814
ECD	AGUR-Net	0.974	0.963	0.947	0.922	0.917
	DAGAN	0.923	0.903	0.896	0.882	0.873
	LSTM-NDT	0.902	0.895	0.883	0.862	0.814
	BiLSTM	0.877	0.871	0.863	0.763	0.741
	MSCDAE	0.891	0.826	0.831	0.846	0.819
	CFlow	0.863	0.825	0.817	0.798	0.774

According to Table 2, on the YDFID-1 dataset, AGUR-Net achieved an F1-score of 0.971 when detecting 50 anomalies, and 0.935 for 250 anomalies. On the ECD dataset, AGUR-Net achieved F1-scores of 0.974, 0.963, and 0.947 for 50, 100, and 150 anomalies, respectively. When detecting 200 anomalies, the F1-scores of LSTM-NDT, BiLSTM, MSCDAE, and CFlow were 0.862, 0.763, 0.846, and 0.798, all lower than AGUR-Net's 0.922. In summary, the AGUR-Net image detection method for anomaly detection in electronic component production can reduce the occurrence of false detections and missed detections and shows good overall performance in inspecting defects in industrial products. To quantify the contribution of each module to the final performance, ablation experiments were designed on the ECD dataset. Using the base U-Net, with all improved modules removed, as the baseline model, SSDAE, MGAN, the improved ECA module, attention gates, and pre-activated residual blocks were gradually added to test anomaly detection performance under each configuration. All experiments were repeated 5 times in the same environment, with different random seeds used to initialize network parameters, and the mean and standard deviation of each indicator were calculated to reflect the stability of the results. The evaluation indicators include AUC, F1 score, Precision, Recall, and single image processing time. The results of the ablation experiment are shown in Table 3.

According to Table 3, after adding SSDAE alone, AUC increased from 0.812 to 0.851, F1 score increased from 0.834 to 0.869, accuracy and recall increased to 0.874 and 0.864, respectively, indicating that denoising processing plays an important role in industrial image anomaly detection. After further adding MGAN, the AUC increased to 0.878, F1 score reached 0.893, accuracy and recall reached 0.896 and 0.890, respectively, verifying the effectiveness of multimodal feature expression for texture-related anomaly detection. Improving the ECA module resulted in a 0.013 increase in AUC and a 0.011 increase in F1 score, as well as a 0.011 and 0.011 increase in accuracy and recall, respectively. At the same time, there was a relatively small increase in parameter count. After adding attention gates and pre-activated residual blocks, the complete method achieved an AUC of 0.922, F1 score of 0.931, accuracy of 0.933, and recall of 0.929, which were improved by 0.110, 0.097, 0.092, and 0.102 respectively compared to the basic U-Net. In terms of the balance between accuracy and recall, the difference between the two indicators of the complete method is only 0.004, indicating that the model has achieved a good balance between reducing false positives and reducing missed detections. In terms of processing time, the complete method takes 378 ms for a single image, which is higher than the 246 ms of the basic U-Net, but still meets real-time requirements in practical industrial detection. The gradual increase of each module did not lead to a significant deterioration in processing time. In addition, the standard deviation of the 5 repeated experiments ranged from 0.005 to 0.012, indicating that the experimental results under all configurations had good stability, and the impact of random initialization on the final performance was relatively small. The above results demonstrate the effectiveness of each proposed module and also indicate that the source of performance improvement in the complete method is the collaborative effect of multiple modules rather than the contribution of a single component.

Table 3 Results of ablation experiment

Model configuration	AUC	F1-score	Precision	Recall	Processing time (ms)
Basic U-Net	0.812 ±0.011	0.834 ± 0.010	0.841± 0.012	0.827 ± 0.009	246
+ SSDAE	0.851± 0.009	0.869 ± 0.008	0.874 ± 0.010	0.864 ± 0.007	283
+ SSDAE + MGAN	0.878 ± 0.008	0.893 ± 0.007	0.896 ± 0.009	0.890 ± 0.008	319
+ SSDAE + MGAN + improved ECA	0.891 ± 0.007	0.904 ± 0.006	0.907 ± 0.008	0.901 ± 0.007	337
Complete method	0.922 ± 0.006	0.931 ± 0.005	0.933 ± 0.007	0.929 ± 0.006	378

4. Discussion

The proposed AGUR-Net anomaly image detection algorithm showed excellent performance in comparison experiments. In terms of anomaly region localization and anomaly image detection accuracy, AGUR-Net outperformed the LSTM-NDT, BiLSTM, and MSCDAE algorithms. In particular, on the MVTecAD and KSDD2 datasets, AGUR-Net maintained high localization accuracy and stability. This performance mainly resulted from its optimized design and strong image feature extraction ability. These results were consistent with the findings of Yan *et al*. and Song *et al*. in 2023 [26, 27]. AGUR-Net also achieved high anomaly image detection accuracy. After 250 iterations, the accuracy of industrial anomaly image detection remained above 75 %, reaching a maximum of 97.26 %, further confirming the algorithm's effectiveness. In the AUC test with a 10 % anomaly rate, AGUR-Net also performed well. Compared with LSTM-NDT, BiLSTM, and MSCDAE, the proposed algorithm consistently achieved an AUC above 0.850, proving its ability to classify industrial products accurately. On the MVTecAD dataset, AGUR-Net reached the highest AUC of 0.946 for the hazelnut samples, while the AUC values of DAGAN, LSTM-NDT, and BiLSTM were 0.832, 0.874, and 0.816 respectively. This advantage was mainly attributed to the improved attention gate module and pre-activation fusion residual block in AGUR-Net, which enhanced its responsiveness to defect regions and its robustness to interference. The results were in agreement with those obtained by Duan *et al*. in their evaluation of anomaly detection algorithms [28].

When the recognition performance for different industrial anomaly images was analyzed, AGUR-Net also demonstrated clear advantages. In the average-loss test for abnormal images from the processing and manufacturing stage of industrial products, the AGUR-Net-based anomaly image detection method for electronic component production achieved an average loss value of 0.016 after 400 iterations. When the number of iterations was between 100 and 200 times, the overall detection average loss value decreased rapidly. In contrast, after 400 iterations, the average loss of LSTM-NDT, BiLSTM, and MSCDAE reached 0.058, 0.061, and 0.067 respectively, all higher than that of the proposed method. This is mainly attributed to the MGAN module in the AGUR-Net-based anomaly image detection method for electronic component production, which improves feature representation and enhances anomaly discrimination in images from the processing and manufacturing stages. Compared with related studies, the anomaly image detection method for electronic component production proposed in this study showed greatly improved performance. In real-time single-image detection speed tests at Factory X2, AGUR-Net also outperformed other methods. It required only 289 ms to process a single image with a type-A defect, while DAGAN, LSTM-NDT, BiLSTM, MSCDAE, and CFlow took 507 ms, 806 ms, 623 ms, 672 ms, and 812 ms respectively. These results proved the superior detection efficiency of AGUR-Net.

The main contributions of this research are twofold: first, a new AGUR-Net anomaly image detection algorithm was designed; second, an image detection method for anomaly detection in electronic component production based on this algorithm was proposed. During the design process, the combination of MGAN and SSDAE networks for image location and classification has improved the accuracy of abnormal target recognition. These contributions have provided new ideas and effective solutions for complex image defect recognition in the fields of intelligent manufacturing and electronic component manufacturing.

5. Conclusion

To address difficulties in identifying defect images and localizing anomalies in the production process of electronic components such as resistors and capacitors, this study proposes an anomaly image detection method for electronic component production based on the AGUR-Net algorithm to improve anomaly detection accuracy in industrial products. The experimental results show that the AGUR-Net abnormal image detection algorithm has high abnormal location accuracy. Furthermore, the empirical results for the proposed anomaly image detection method for electronic component production show that, in single-image processing tests on industrial products, the method requires less time and achieves higher overall detection efficiency. Moreover, anomaly localization is carried out on images from the stages of raw material preparation, processing and manufacturing, surface treatment, quality inspection, and packaging and transportation of parts. The AGUR-Net anomaly image detection method for electronic component production has a better anomaly localization effect on images in the surface treatment stage, and the heat-map distribution area is relatively concentrated. Overall, the AGUR-Net-based anomaly image detection method for electronic component production can not only accurately classify industrial targets, but also precisely localize anomalies in the production processes of different industrial products. However, there are still limitations to the research, and its generalization ability in factory production lines or other industrial scenarios still needs further testing. There may be significant differences between different factories in terms of image acquisition equipment, lighting conditions, product materials and textures, defect type distribution, etc., all of which may lead to a decrease in model performance. Therefore, future work can proceed in the following directions: first, conducting cross-factory transfer learning experiments to verify the degree of performance degradation when the model is directly applied to a target-domain factory after training in the source domain. Second, exploring few-sample or unsupervised domain adaptation methods to enable the model to adapt quickly using only a small number of unlabeled images from the target factory. Finally, the prototype system should be deployed in the actual production line and a long-term online evaluation should be conducted, collecting real testing data from different batches, different lighting conditions, and different product specifications to comprehensively measure the robustness and deployability of the method.

Acknowledgement

This work was supported by the Natural Science Foundation of Chengdu Aeronautic Polytechnic University (Grant No. ZZX0623061) and the Indoor Space Layout Optimization and Safety Assurance Laboratory of Chengdu Normal University (Project No. SNKJ202513).

References

[1] Leung, J.-H., Tsao, Y.-M., Karmakar, R., Mukundan, A., Lu, S.-C., Huang, S.-Y., Saenprasarn, P., Lo, C.-H., Wang, H.-C. (2024). Water pollution classification and detection by hyperspectral imaging, *Optics Express*, Vol. 32, No. 14, 23956-23965, doi: 10.1364/OE.522932.

[2] Zhao, T., Gu, Y., Yang, J., Usuyama, N., Lee, H.H., Kiblawi, S., Naumann, T., Gao, J., Crabtree, A., Abel, J., Moung-Wen, C., Piening, B., Bifulco, C., Wei, M., Poon, H., Wang, S. (2025). A foundation model for joint segmentation, detection and recognition of biomedical objects across nine modalities, *Nature Methods*, Vol. 22, No. 1, 166-176, doi: 10.1038/s41592-024-02499-w.

[3] Abid, A., Jemili, F., Korbaa, O. (2024). Real-time data fusion for intrusion detection in industrial control systems based on cloud computing and big data techniques, *Cluster Computing*, Vol. 27, No. 2, 2217-2238, doi: 10.1007/s10586-023-04087-7.

[4] Yang, Z., Zhang, M., Chen, Y., Hu, N., Gao, L., Liu, L., Song, J.I. (2024). Surface defect detection method for air rudder based on positive samples, *Journal of Intelligent Manufacturing*, Vol. 35, No. 1, 95-113, doi: 10.1007/s10845-022-02034-8.

[5] Ma, Y., Yin, J., Huang, F., Li, Q. (2024). Surface defect inspection of industrial products with object detection deep networks: A systematic review, *Artificial Intelligence Review*, Vol. 57, Article No. 333, doi: 10.1007/s10462-024-10956-3.

[6] Ullah, W., Khan, S.U., Kim, M.J., Hussain, A., Munsif, M., Lee, M.Y., Baik, S.W. (2024). Industrial defective chips detection using deep convolutional neural network with inverse feature matching mechanism, *Journal of Computational Design and Engineering*, Vol. 11, No. 3, 326-336, doi: 10.1093/jcde/qwae019.

[7] Liu, R., Xiao, D., Lin, D., Zhang, W. (2024). Intelligent bearing anomaly detection for industrial Internet of Things based on auto-encoder Wasserstein generative adversarial network, *IEEE Internet of Things Journal*, Vol. 11, No. 13, 22869-22879, doi: 10.1109/JIOT.2024.3358871.
[8] Gong, Y., Wang, X., Zhou, C., Ge, M., Liu, C., Zhang, X. (2025). Human-machine knowledge hybrid augmentation method for surface defect detection based few-data learning, *Journal of Intelligent Manufacturing*, Vol. 36, No. 3, 1723-1742, doi: 10.1007/s10845-023-02270-6.
[9] Bergmann, P., Batzner, K., Fauser, M., Sattlegger, D., Steger, C. (2022). Beyond dents and scratches: Logical constraints in unsupervised anomaly detection and localization, *International Journal of Computer Vision*, Vol. 130, 947-969, doi: 10.1007/s11263-022-01578-9.
[10] Liu, Y., Zhang, Y., Jiang, Y., Liu, W., Yang, F. (2023). UWB-INS fusion positioning based on a two-stage optimization algorithm, *Tehnički Vjesnik – Technical Gazette*, Vol. 30, No. 1, 185-190, doi: 10.17559/TV-20221019035741.
[11] Zhang, H., Guo, W., Zhang, S., Lu, H., Zhao, X. (2022). Unsupervised deep anomaly detection for medical images using an improved adversarial autoencoder, *Journal of Digital Imaging*, Vol. 35, No. 2, 153-161, doi: 10.1007/s10278-021-00558-8.
[12] Sun, Z., Wang, J., Li, Y. (2024). RAMFAE: A novel unsupervised visual anomaly detection method based on auto-encoder, *International Journal of Machine Learning and Cybernetics*, Vol. 15, 355-369, doi: 10.1007/s13042-023-01913-7.
[13] Wei, Z.H., Yan, L., Yan, X. (2024). Optimizing production with deep reinforcement learning, *International Journal of Simulation Modelling*, Vol. 23, No. 4, 692-703, doi: 10.2507/IJSIMM23-4-CO17.
[14] Zhang, H., Wang, S., Lu, S., Yao, L., Hu, Y. (2023). Attention-Gate-based U-shaped Reconstruction Network (AGUR-Net) for color-patterned fabric defect detection, *Textile Research Journal*, Vol. 93, No. 15-16, 3459-3477, doi: 10.1177/00405175221149450.
[15] Li, C., Qi, L., Geng, X. (2025). A SAM-guided two-stream lightweight model for anomaly detection, *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol. 21, No. 2, 1-23, doi: 10.1145/3706574.
[16] Korovin, A., Vasilev, A., Egorov, F., Saykin, D., Terukov, E., Shakhray, I., Zhukov, L. Budennyy, S. (2023). Anomaly detection in electroluminescence images of heterojunction solar cells, *Solar Energy*, Vol. 259, 130-136, doi: 10.1016/j.solener.2023.04.059.
[17] Yin, S., Li, H., Laghari, A.A., Gadekallu, T.R., Sampedro, G.A., Almadhor, A. (2024). An anomaly detection model based on deep auto-encoder and capsule graph convolution via Sparrow Search Algorithm in 6G Internet of Everything, *IEEE Internet of Things Journal*, Vol. 11, No. 18, 29402-29411, doi: 10.1109/JIOT.2024.3353337.
[18] Abdelmawla, A., Ma, S., Yang, J.J., Kim, S.S. (2023). Subsurface anomaly detection utilizing synthetic GPR images and deep learning model, *Geomechanics and Engineering*, Vol. 33, No. 2, 203-209, doi: 10.12989/gae.2023.33.2.203.
[19] Liu, H., Li, L. (2023). Anomaly detection of high-frequency sensing data in transportation infrastructure monitoring system based on fine-tuned model, *IEEE Sensors Journal*, Vol. 23, No. 8, 8630-8638, doi: 10.1109/JSEN.2023.3254506.
[20] Faber, K., Corizzo, R., Sniezynski, B., Japkowicz, N. (2023). VLAD: Task-agnostic VAE-based lifelong anomaly detection, *Neural Networks*, Vol. 165, 248-273, doi: 10.1016/j.neunet.2023.05.032.
[21] Yun, H., Kim, H., Jeong, Y.H., Jun, M.B.G.(2023). Autoencoder-based anomaly detection of industrial robot arm using stethoscope based internal sound sensor, *Journal of Intelligent Manufacturing*, Vol. 34, 1427-1444, doi: 10.1007/s10845-021-01862-4.
[22] Wang, Y., Fang, R. (2023). An approach for fast fault detection in virtual network, *Tehnički Vjesnik – Technical Gazette*, Vol. 30, No. 4, 1146-1151, doi: 10.17559/TV-20230207000330.
[23] Wang, X., Liu, C., Hou, J., Zhou, L. (2024). Advancing crack segmentation detection: Introducing AAMC-Net algorithm for image crack analysis, *Computer Science and Information Systems*, Vol. 21, No. 4, 1435-1455, doi: 10.2298/CSIS230725042W.
[24] Liu, J., Zhou, G. (2024). Learning discriminative representations through an attention mechanism for image-based person re-identification, *Computer Science and Information Systems*, Vol. 21, No. 4, 1483-1498, doi: 10.2298/CSIS230829044L.
[25] Aka, A.C., Atta, A.F., Keupondjo, S.G.A., Oumtanaga, S. (2023). An efficient anchor-free localization algorithm for all cluster topologies in a wireless sensor network, *International Journal of Computers Communications & Control*, Vol. 18, No. 3, Article No. 4961, doi: 10.15837/ijccc.2023.3.4961.
[26] Yan, S., Chen, P., Chen, H., Mao, H., Chen, F., Lin, Z. (2024). Multiresolution feature guidance based transformer for anomaly detection, *Applied Intelligence*, Vol. 54, No. 2, 1831-1846, doi: 10.1007/s10489-024-05283-7.
[27] Song, J., Lee, Y.C., Lee, J. (2023). Deep generative model with time series-image encoding for manufacturing fault detection in die casting process, *Journal of Intelligent Manufacturing*, Vol. 34, No. 7, 3001-3014, doi: 10.1007/s10845-022-01981-6.
[28] Duan, M., Mao, L., Liu, R., Liu, W., Liu, Z. (2024). Unified model based on reinforced feature reconstruction for metro track anomaly detection, *IEEE Sensors Journal*, Vol. 24, No. 4, 5025-5038, doi: 10.1109/JSEN.2023.3348118.