

Chest X-ray Abnormality Detection Using Convolutional AutoEncoder Combined with Double Generative Adversarial Network (GAN)

Tan Yanli^{1,2}, Azliza Mohd Ali^{2*}, Sharifalillah Nordin², Wang Jin^{1,2}, and Li Guoqin^{1,2}

¹Department of Electronic Engineering, Taiyuan Institute of Technology, Taiyuan, Shanxi, China; tanyanli@studysedu.cn

²College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Malaysia

*Correspondence: Azliza Mohd Ali; azliza@tmsk.uitm.edu.my; Tel.: +60198726432

ABSTRACT- The statistical properties of aberrant samples are unstable, and traditional chest X-ray image data is difficult to gather and unevenly distributed. A CAE-D-GAN anomaly detection model based on dual GANs and convolutional autoencoders is proposed in this paper. We employ the DGAN module to obtain clear degraded images, the convolutional autoencoder to extract low-dimensional features from high-dimensional data, and the DDGAN module to learn how to degrade images and recover clearer images from degraded photos during the adversarial process. While the reconstruction error score identifies sample flaws, the network is optimized by the error loss between the original and reconstructed samples. The network just needs normal samples to be trained, and it can reach a maximum AUROC value of 0.86. The findings demonstrate that the CAE-D-GAN model outperforms a number of different anomaly detection models in terms of detection effects and feature reconstruction capabilities. There are special opportunities for using this approach to detect other anomalies in medical images.

Keywords: Abnormal detection, Autoencoder, CAE-D-GAN, Reconstruction loss, Anomaly score.

ARTICLE INFORMATION

Author(s): Tan Yanli, Azliza Mohd Ali, Sharifalillah Nordin, Wang Jin, and Li Guoqin;

Received: 10/12/2024; **Accepted:** 03/08/2025; **Published:** 30/09/2025;

E- ISSN: 2347-470X;

Paper Id: IJEER 1012-05;

Citation: 10.37391/ijeer.130321

Webpage-link:

<https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130321.html>



Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

It is challenging to identify abnormalities in medical imaging because abnormal samples are rare and challenging to obtain, and the distribution of lesion data is uneven and varied. As a result, we propose a method that uses dual GANs and convolutional autoencoders to automatically detect anomalies in chest X-ray pictures. The convolutional autoencoder has a lower reconstruction error since it learns reconstruction features from normal data [1]. The network design of the generative adversarial network is founded on an adversarial mechanism. Both the discriminative and generative models keep playing games of minimization and maximization while they are being trained. As a result, the model has a strong capacity to produce data [2]. Compared to the conventional unsupervised learning framework, the generative model, which collaborates with the discriminative model, requires less previous knowledge [3,4]. It can also efficiently learn the distribution of actual data and modify its settings to get the best outcomes. Convolutional autoencoder, DGAN, and DDGAN are the three modules that make up the suggested model. We compress the

original data from the high-dimensional data space into feature representations in the low-dimensional data space and utilize the convolutional autoencoder module to learn low-dimensional features from high-dimensional data. To acquire paired, clear, degraded images and train on unpaired datasets, we primarily utilize the DGAN module. Realistic, deteriorated images are produced by DGAN. One module that learns to deteriorate is the DDGAN module. During the adversarial process, it recovers sharper images from deteriorated ones in order to identify irregularities.

2. ALGORITHM MODEL STRUCTURE

2.1. CAE-D-GAN Model Structure

As you can see in *figure 1*, we developed a novel method for automatically detecting anomalies called CAE-D-GAN, which consists of a double GAN and a convolutional autoencoder. The three modules that make up the model are the DGAN, DDGAN, and convolutional autoencoder modules. We acquire paired, clear, degraded images using the DGAN module, learn to degrade using the DDGAN module, and extract low-dimensional features from high-dimensional data using the convolutional autoencoder module. After training the DGAN on unpaired datasets, the discriminator oversees it to generate more realistic damaged images. In order to create clear images, the DDGAN generating network trains on paired, clear, degraded images after receiving the realistic degraded images from DGAN. The generated clear images and genuine images are sent to the discriminator network of DDGAN, which plays a game with the generator and faces off against it. In order to identify irregularities, the confronting procedure recovers sharper images from deteriorated ones.

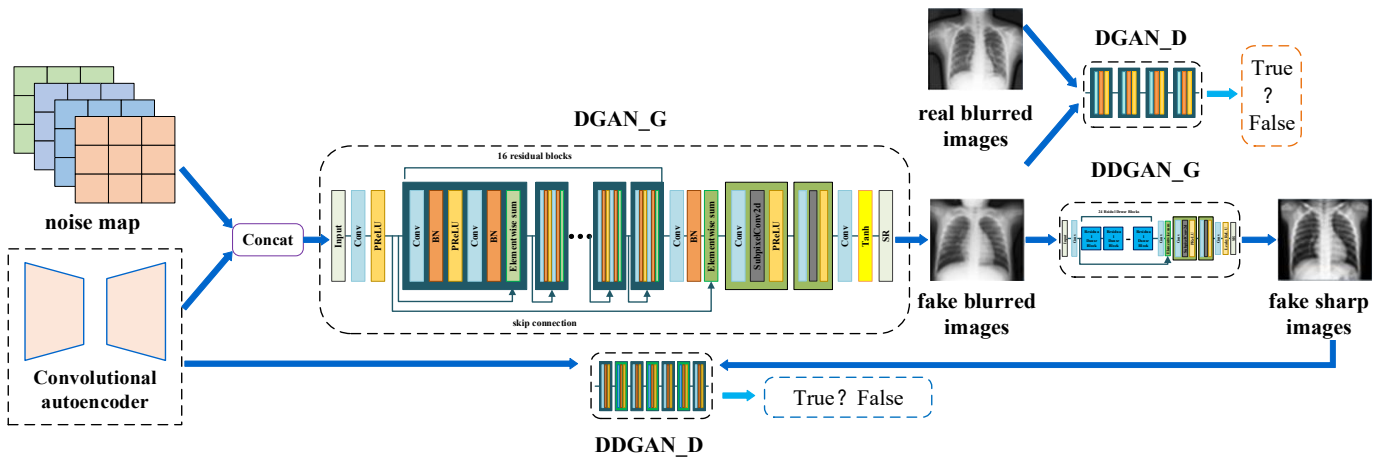


Figure 1. CAE-D-GAN Model Architecture

A restructured power system is one that moves away from the old, vertically integrated utility model and towards a more decentralised, competitive, and market-based structure. Deregulation in the electrical sector refers to the progression of amending the law and policy governing corporate operations to provide consumers with a selection of energy firms. In a deregulated electricity market as shown in figure 1, consumers and energy suppliers have the freedom to invest their resources in the preservation and perfection of power production and transmission infrastructure as they perceive vigorous. The power generated by Genco's assets is distributed extensively via the trade utilities network. Customers get advantages through the ability to assess and contrast prices and amenities, along with the adoption of stochastic techniques.

2.2. Algorithm Framework

Part of the pseudocode may be seen in figure 2. The input image is down-sampled and its dimensionality reduced to obtain the encoded representation of the hidden layer. Up-sampling and dimensional augmentation are then used to reconstruct the input data for the DGAN generator. The discriminator oversees the creation of an authentically damaged image. A link to the DDGAN network is made in series. This dual GAN series connection requires the model to preserve information necessary for reconstructing data samples. To reduce the overall reconstruction error, the information that is kept must be as pertinent to most typical samples as feasible. As a result, data samples that are not representative of the main dataset are not adequately rebuilt. The pixel differences between the input and reconstructed images can be used to learn more about anomaly detection. Variations in the reconstruction error's distribution indicate anomalies, and the data reconstruction error serves as an anomaly score. During the training phase, the network weights are optimized using the reconstruction error. During the inference stage, backpropagation is used to iteratively determine the feature representation of the test image's hidden space.

```
def main_test(self,option=1):
    values, labels=self.get_roc(option)
    TPRs=[]
    FPRs=[]
```

```
best_acc=0
best_idx=0
for i in tqdm(np.arange(0,1,0.01)):
    TPR,FPR,
    acc=self.split_values(values,labels,i)
    FPRs.append(FPR)
    TPRs.append(TPR)
    if acc>best_acc:
        best_acc=acc
        best_idx=i
```

Figure 2. A portion of the pseudocode

2.3. Autoencoder Model Architecture

The convolutional encoder, convolutional decoder, and bottleneck module that connects them make up the convolutional autoencoder's network topology, which is depicted in figure 3. In order to increase the dimension and reconstruct the input data, the convolutional encoder downsamples the input data [5], and the convolutional decoder upsamples the encoded representation of the hidden layer to decrease the dimension [6].

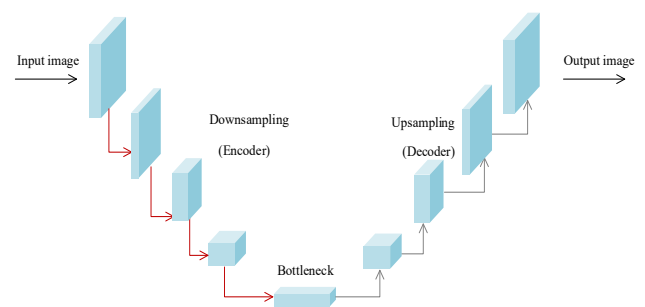


Figure 3. Convolutional autoencoder network structure

2.3. DGAN Network Structure

2.3.1 DGAN Generator Network

The generator network's structure is shown in Figure 4. An input component, a residual network portion made up of 16 residual networks, an image magnification part, and an output part make up the DGAN generation network.

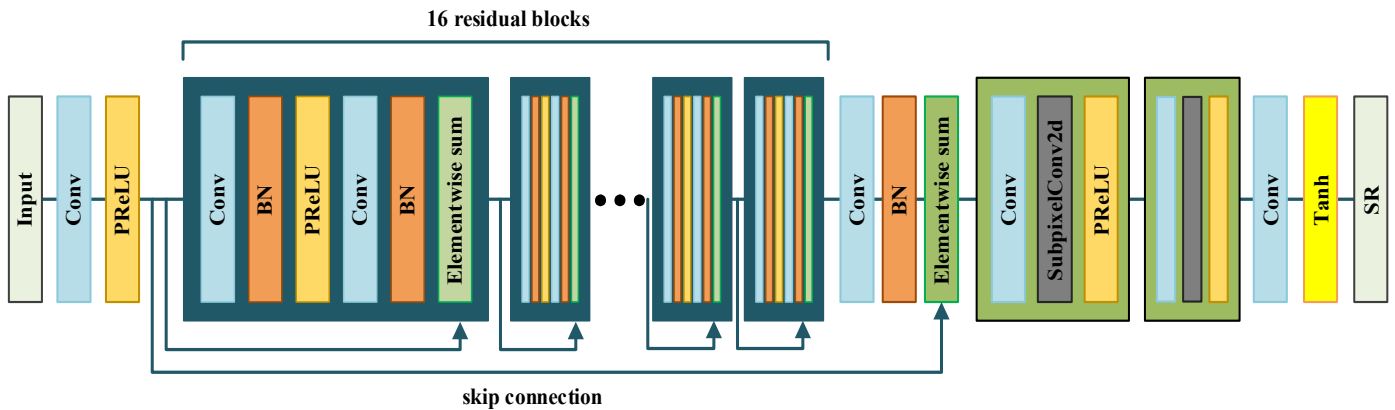


Figure 4. DGAN Generator Network Structure

A residual network comes after the image has gone through a convolution layer with a 3*3 kernel and an activation function layer. There are sixteen residual blocks in the residual network [7]. To create a deep convolutional network, we add a skip connection between each of the two residual blocks. This improves the model's gradient stability throughout training.

2.3.2. DGAN Discriminator Network

The discriminator network structure of the DGAN is shown in *figure 5*. It is composed of three parts: input, convolution, and output. Different convolution kernels used in the convolution layer of the convolution component are distinguished by light blue and dark blue, respectively. Using the output part's elementwise sum layer, we fuse features [8]. By keeping all extracted feature variables, regularization methods lower the order of magnitude of the variables.

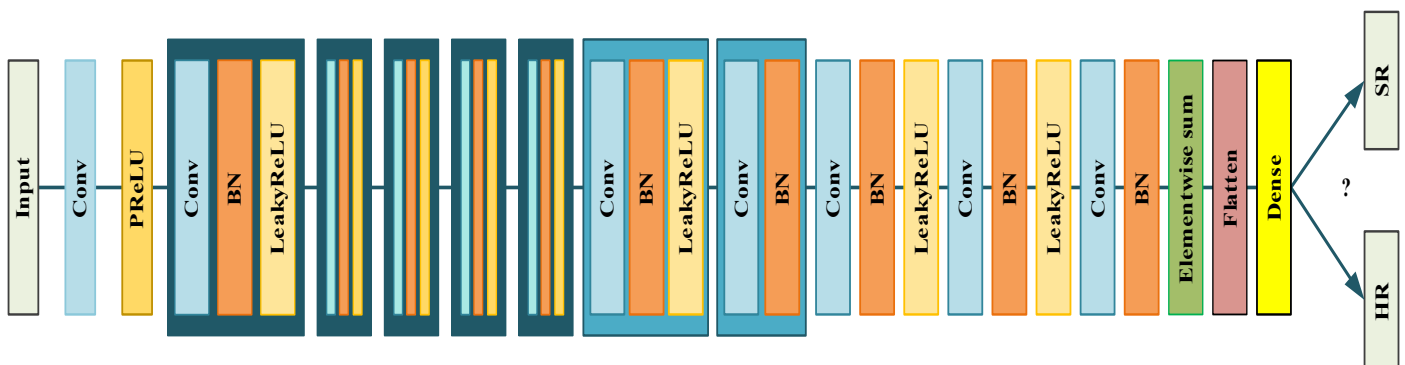


Figure 5. DGAN Discriminant Network Structure

2.4. DDGAN Network Structure

2.4.1 DDGAN Generator Network

The DDGAN's generative network structure is displayed in *figure 6*. In DDGAN's generative network, we have an input part, an RDB part with 24 residual dense blocks that combine the DenseNet structure and residual network, an image magnification part, and an output part. The RDB component integrates the DenseNet network structure with the concept of residual networks. Each RDB is made up of three DenseNet blocks that are tightly coupled via residual scaling. The network can produce a lot of features with just a few convolution kernels.

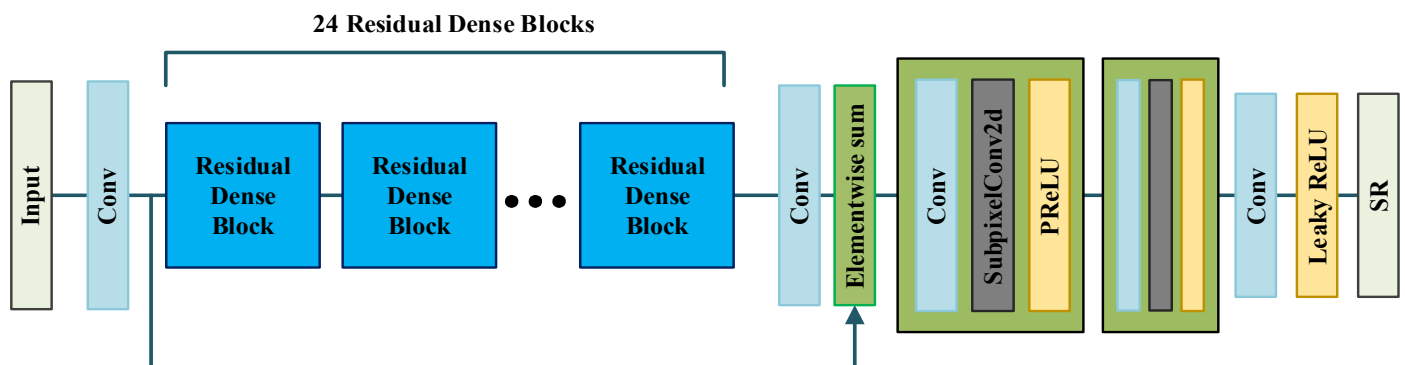


Figure 6. DDGAN Generation Network Structure

2.4.2. DDGAN Discriminator Network

The DDGAN's discriminant network is displayed in *figure 7*. An input portion, a convolution part, and an output part make up the DDGAN discriminant network. Using dark blue and green modules, the convolution layer of the convolution section differentiates between various convolution kernels.

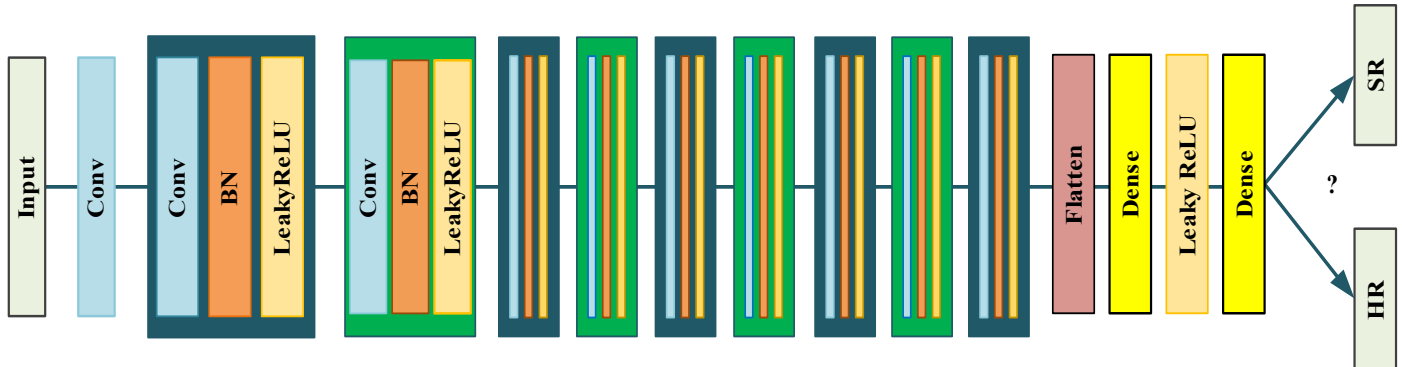


Figure 7. DDGAN Discriminant Network Structure

3. MODEL TRAINING

To optimize the model parameters during the training process and improve the reconstruction ability of normal samples, the following three loss functions are set: generator loss function L_G , discriminator loss function L_D , and pixel-level loss function L_P [9].

3.1. Generator loss function L_G

We propose the generator loss function to enhance the quality of the reconstructed image and generate a reconstructed output that closely resembles the original input data. The generator loss function L_G includes three sub-functions: reconstruction loss L_{recon} , encoding loss L_{en} , and adversarial loss L_{adv} .

The reconstruction loss is the difference per pixel between the input image and the reconstructed image [10], as defined below:

$$L_{recon} = E_{x \sim p_r(x)} \|x - G(x)\|_1 \quad (1)$$

Among them, x is the input image, $G(x)$ is the reconstructed output image, and $p_r(x)$ is the input normal data distribution.

Coding loss L_{en} is used to measure the similarity between the feature coding z of the input image and the feature coding $\hat{z}$ of the reconstructed output, as defined below:

$$L_{en} = E_{x \sim p_r(x)} \|G_{en}(x) - G_{en}(G(x))\|_1 \quad (2)$$

Among them, $G_{en}(x)$ represents the feature encoding of the input image.

The generator's adversarial loss L_{adv} aims to replicate the original image as closely as possible in the reconstructed image. For the generator G , it is necessary to minimize the adversarial loss; for the discriminator D , it is necessary to maximize the adversarial loss. Through such a game, it is finally achieved $\min_G \max_D L_{adv}$, as defined below:

$$L_{adv} = -E_{x \sim p_r(x)} [D(G(x))] \quad (3)$$

The total generator loss function is expressed as follows:

$$L_G = \lambda_{recon} L_{recon} + \lambda_{en} L_{en} + \lambda_{adv} L_{adv} \quad (4)$$

Among them, λ_{recon} , λ_{en} , λ_{adv} represent the weight parameters, respectively.

3.2. Discriminator loss function L_D

The concept of a binary cross-entropy loss function is used by the discriminator loss function [11]. The discriminator can separate created data from real data as much as feasible by reducing this disparity. The following is its definition:

$$L_D = E_{x \sim p_r(x)} [\max(0, 1 - D(x))] + E_{\hat{x} \sim p_g(x)} [\max(0, 1 + D(\hat{x}))] \quad (5)$$

3.3. Pixel-level loss function L_P

To confine the generator's training in the image reconstruction task, the pixel-level distance between the rebuilt and original images must be determined [12]. The following is its definition:

$$L_P = \frac{1}{N} \sum_i^N \|\hat{x} - x\|_1 \quad (6)$$

From the standpoints of limiting feature semantics, constraining reconstruction errors [13], and adversarial learning using three distinct loss functions, the reconstruction generator enhances the model's capacity to generate reconstructions.

4. EXPERIMENTAL SETUP

4.1. Dataset

With a primary focus on three image types—normal lungs, COVID-19, and viral pneumonia—the datasets used in this work are taken from three public datasets (<http://www.kaggle.com/tawsifurrahman/covid-19radiography-database>, <https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia>, <https://www.kaggle.com/pranavraikokte/covid19->

image-dataset). There are 4,961 aberrant chest X-ray photos and 10,192 normal chest X-ray images in dataset 1. There are 4,849 aberrant chest X-ray photos and 1,583 normal chest X-ray images in dataset 2. There are 222 abnormal chest X-ray photos and 90 normal chest X-ray images in dataset 3.

Every set of data is separated into training and test sets. Half of the normal images are chosen at random to serve as the training set, and the remaining normal images are combined with abnormal images in four different proportions to create the test set: 25%, 50%, 75%, and 100%.

4.2. Model Parameters

During the training process, the Adam [14] optimizer is used, the learning rate is set to $1e^{-4}$, the weight penalty coefficient is set to 0, the momentum coefficient $\beta_1 = 0.9$, $\beta_2 = 0.999$, and the weight of the loss function are set to $\lambda_{adv} = 1$, $\lambda_{recon} = 50$, $\lambda_{en} = 1$, all models are iterated over the entire training set (epoch) and trained 10, 20, and 50 times, each time before The number of samples (batch size) for forward propagation and backpropagation is 159. We implement the model using the PyTorch framework version 1.5.1, and train it on an NVIDIA GeForce GTX 1080 GPU.

4.3. Evaluation indicators

The samples are classified as follows: false negative (FN: abnormal sample is considered normal), true negative (TN: normal sample is correctly detected), false positive (FP: normal samples are considered abnormal), and true positive (TP: abnormal sample is successfully detected) [15]. In light of this, we define the true positive rate (TPR) and false positive rate (FPR) and use precision rate (P), recall rate (R), and F1 score as assessment indicators.

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (9)$$

$$FPR = \frac{FP}{FP+TN} \quad (10)$$

$$TPR = \frac{TP}{TP+FN} \quad (11)$$

To get the ROC curve, connect the coordinate points below the threshold, using TPR as the ordinate and FPR as the abscissa. The performance index, or AUROC, is the area under the curve [16]. The network's ability to detect anomalies improves with a higher AUROC rating.

In order to assess the quality of the network-reconstructed image, the study additionally established two assessment indicators: SSIM[17] and PSNR[18]. The difference in structural information between the original image and the reconstructed image can be more accurately reflected by SSIM. By comparing the mean square error between the original and rebuilt images, PSNR calculates the quality of the reconstructed picture[19]. The PSNR value increases and the quality of the reconstructed

image improves with a lower difference between the corresponding pixels of the two images. The following is its definition:

$$SSIM(x, \hat{x}) = \frac{(2\mu_x\mu_{\hat{x}}+c_1)(2\sigma_{x\hat{x}}+c_2)}{(\mu_x^2+\mu_{\hat{x}}^2+c_1)(\sigma_x^2+\sigma_{\hat{x}}^2+c_2)} \quad (12)$$

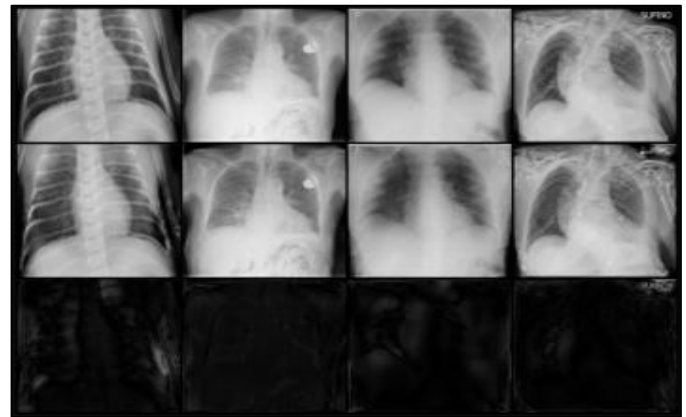
Where μ_x and $\mu_{\hat{x}}$ are the average values of images x and $\hat{x}$, σ_x^2 and $\sigma_{\hat{x}}^2$ are the variances of images x and $\hat{x}$, and $\sigma_{x\hat{x}}$ is the covariance of images x and $\hat{x}$.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_x^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_x}{\sqrt{MSE}} \right) \quad (13)$$

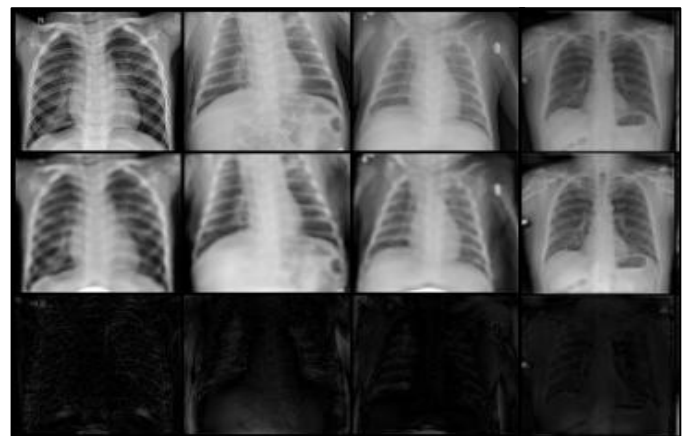
Among them, MAX_x is the maximum possible pixel value of the image, and MSE is the mean square error of the input image.

5. EXPERIMENTAL RESULTS

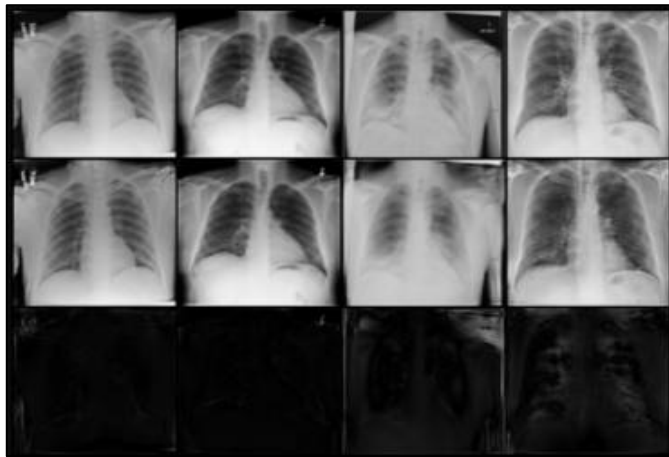
The application of the suggested model CAE-D-GAN to the reconstruction impact of chest X-ray datasets 1, 2, and 3 is demonstrated in figure 8. The input chest X-ray test image is shown in the first row, followed by the chest X-ray image that was reconstructed using the suggested network in the second row, and the difference between the input and reconstructed values in the third row. We can see that aberrant images are successfully and with little reconstruction error reconstructed by the suggested network.



dataset1



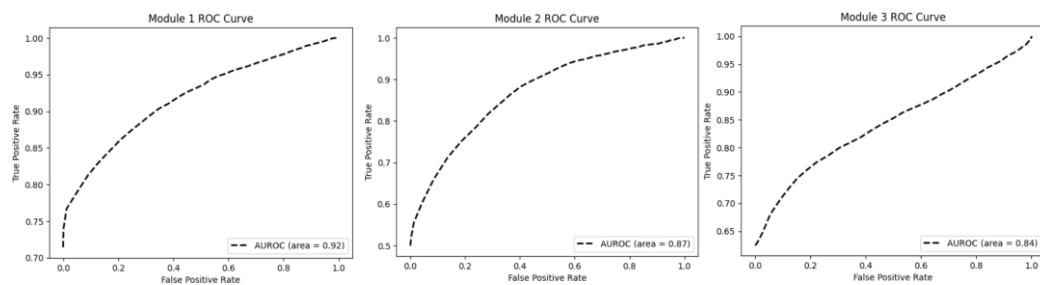
dataset2



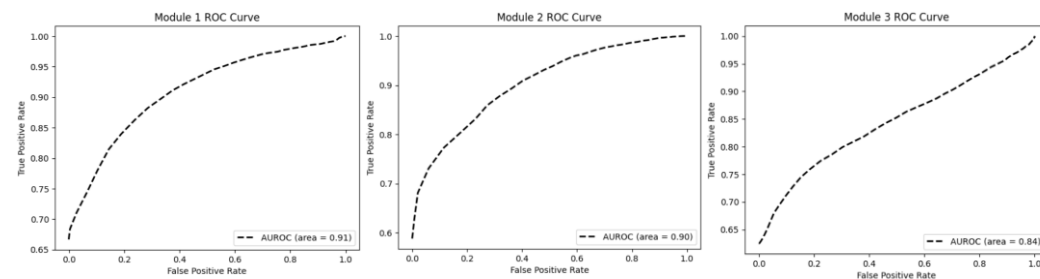
dataset3

Figure 8. Reconstruction Effect on dataset1, dataset2, and dataset3 based on the CAE-D-GAN Model

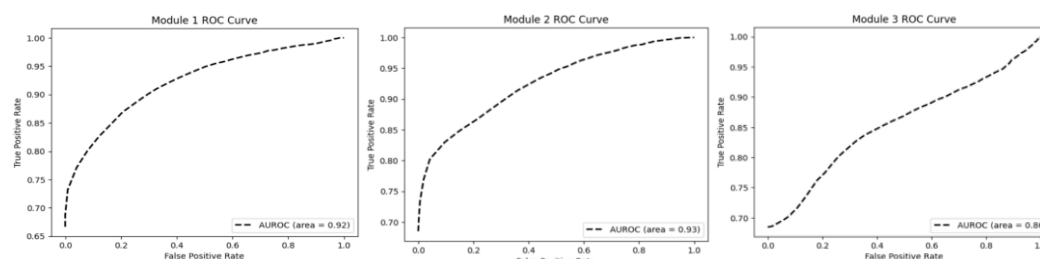
AUROC primarily determines if an image is normal or abnormal based on the image as a whole. With Module 1 denoting the convolutional autoencoder module, Module 2 denoting the DGAN module, and Module 3 denoting the DDGAN module, *figure 9* displays the AUROC evaluation index findings of the proposed CAE-D-GAN model on dataset 1 for tests of varying proportions. Following testing, we discover that the maximum AUROC values for Modules 1, 2 and 3 at a 25% ratio are 0.92, 0.87, and 0.84, respectively. Module 1's maximum AUROC value is 0.91, Module 2's maximum AUROC value is 0.90, and Module 3's maximum AUROC value is 0.84 at a 50% ratio. Module 1's maximum AUROC value is 0.92, Module 2's maximum AUROC value is 0.93, and Module 3's highest AUROC value is 0.86 at a 75% ratio. Module 1's maximum AUROC value is 0.91, Module 2's maximum AUROC value is 0.90, and Module 3's maximum AUROC value is 0.85 at a 100% ratio.



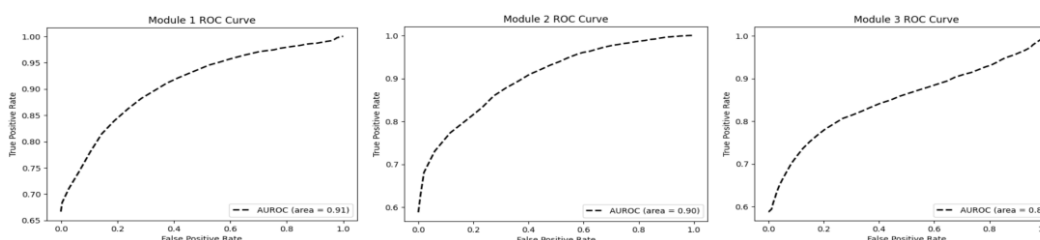
(a) AUROC of the Test Set of Dataset1 at a Ratio of 25%



(b) AUROC of the Test Set of Dataset1 at a Ratio of 50%



(c) AUROC of the Test Set of Dataset1 at a Ratio of 75%



(d) AUROC of the Test Set of Dataset1 at a Ratio of 100%

Figure 9. AUROC Results of Dataset 1 in Different Test Sets

The AUROC evaluation index values for test sets of varying proportions that the algorithm in this paper produced are displayed in *figure 10*. The model can get an AUROC value of up to 0.86. The algorithm's employment of two GANs is the cause. While the second GAN learns to degrade, the algorithm uses the first GAN to produce clear degraded images. To improve anomaly detection findings and recover clearer images from degraded photographs, we train it on paired clear and degraded images.

The F1-Confidence curve after 50 epochs is shown in *figure 11*. The graphic shows that the F1 value for all chest X-ray image recognition at a confidence level of 0.82 is centered at 0.272, indicating robust model recognition performance and a respectable balance between precision and recall. The model has a good recognition accuracy and predicts positive values.

Table 1 displays the acquired precision, recall, F1 score, PSNR, and SSIM values. The PSNR value is approximately 28.3, which is within a reasonably consistent range, as the number of epochs grows. The model produces good picture quality, and the reconstructed image quality is high, as indicated by the SSIM values, which are primarily around 0.93. The loss function value of dataset 1 is shown as a curve in *figure 12* as the iteration procedure varies. As the loss gap between the rebuilt and original images narrows and eventually stabilizes, it is evident that the network has a good reconstruction capacity for normal samples.

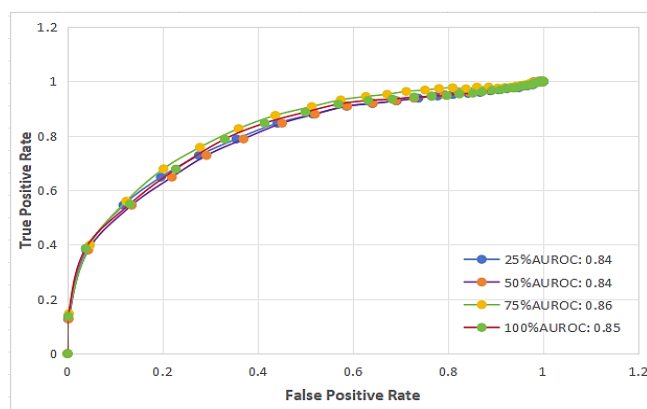


Figure 10. Comparison of the AUROC curves obtained with different sample proportions

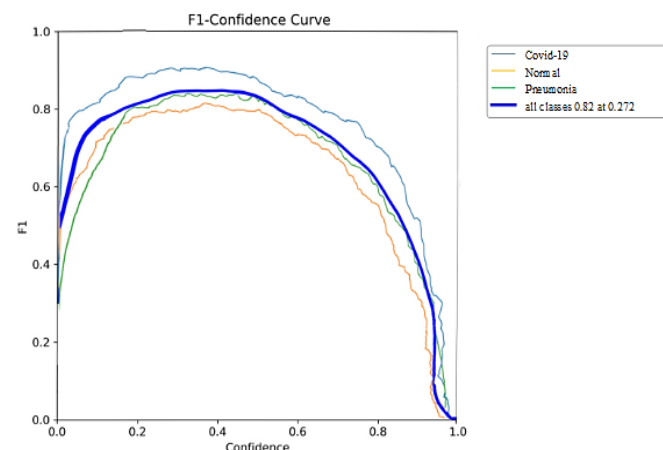


Figure 11. F1-Confidence Curve

Table 1. Performance index values obtained for different data sets with epochs of 10, 20, and 50

Epoch	25%	50%	75%	100%	Parameters
10	0.80	0.82	0.83	0.85	Precision
	0.95	0.93	0.90	0.89	Recall
	0.88	0.86	0.89	0.83	F1-score
	26.35	28.17	28.63	28.46	PSNR
	0.92	0.93	0.93	0.93	SSIM
20	0.82	0.80	0.85	0.84	Precision
	0.95	0.92	0.93	0.92	Recall
	0.90	0.89	0.87	0.85	F1-score
	27.23	28.59	28.66	28.85	PSNR
	0.92	0.94	0.92	0.93	SSIM
50	0.81	0.83	0.85	0.84	Precision
	0.93	0.94	0.91	0.92	Recall
	0.88	0.91	0.89	0.88	F1-score
	27.56	28.68	28.49	28.27	PSNR
	0.92	0.93	0.93	0.94	SSIM

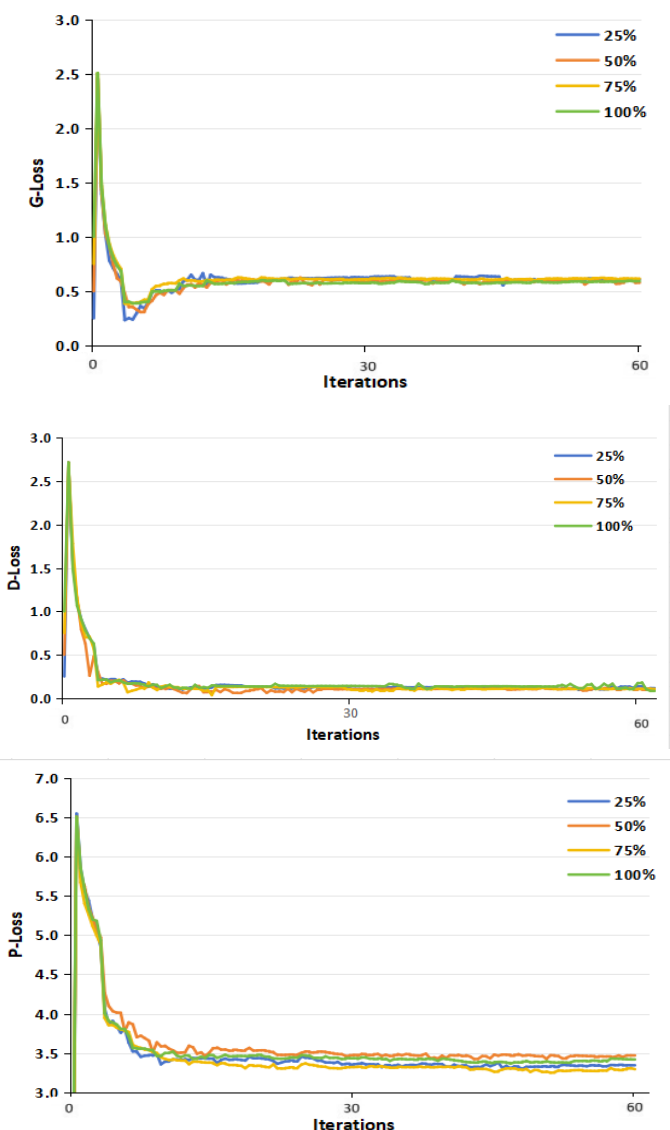


Figure 12. G-Loss, D-Loss, and P-Loss over training iterations under different scale data sets

The detection outcomes of the CAE-D-GAN method are compared in this study with those of GANomaly [20], Efficient-GAN [21], Skip-GAN [22], and EGBAD [23]. The suggested approach is algorithmically based on an unsupervised learning approach and is structurally made up of a dual GAN cascade and a convolutional autoencoder. The suggested CAE-D-GAN model differs from models like GANomaly mostly in these two areas. The training parameter setups for every model are shown in *table 2*. The five models' best AUROC values on chest X-ray datasets at various sizes are shown in *figure 13*. The CAE-D-GAN model, which performs best in detecting abnormalities in chest X-rays, is proposed in this work. The performance indicators of the five models on the chest X-ray dataset are shown in *table 3*. It can be seen that the proposed model has high detection accuracy, good image reconstruction ability and high image signal-to-noise ratio. The proposed approach has clear benefits in terms of interpretability, training speed, and reconstruction quality.

Table 2. The training parameter settings

Methods	Input sizes	Batch sizes	Epochs	Whether each model was re-trained
GANomaly	256×256	64	50	Y
Efficient-GAN	256×256	64	50	Y
Skip-GAN	128×128	64	50	Y
EGBAD	128×128	64	50	Y
CAE-D-GAN	128×128	64	50	Y

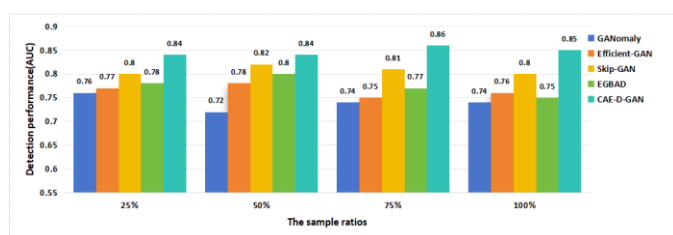


Figure 13. The detection accuracy of different models under different sample proportions

Table 3. Five Unsupervised Anomaly Detection Model Performance Metrics

Parameters Methods	Precision	Recall	F1-score	AUROC	PSNR	SSIM
GANomaly	0.80	0.90	0.84	0.76	26.19	0.79
Efficient-GAN	0.81	0.86	0.86	0.78	26.58	0.83
Skip-GAN	0.83	0.89	0.90	0.82	27.93	0.91
EGBAD	0.83	0.93	0.88	0.80	27.32	0.89
CAE-D-GAN	0.85	0.95	0.91	0.86	28.46	0.94

6. DISCUSSION AND CONCLUSIONS

To overcome the imbalance of positive and negative samples in chest X-ray pictures, we develop an unsupervised anomaly detection technique based on normal samples. The technique extracts the possible feature vector distribution in normal data more precisely by utilizing a network structure that combines a double GAN cascade with a convolutional autoencoder. In order to rebuild high-quality photos, we employ the GAN module to learn the de-degradation ability during the game process. In order to reach convergence, we improve the network model using different loss functions, including SSIM. The CAE-D-GAN model presented in this paper outperforms the earlier advanced GAN anomaly detection model for the chest X-ray dataset. By using only normal samples for training, we are able to detect chest X-ray anomalies.

The main contributions of this paper are:

1. For the first time, we apply the unsupervised anomaly detection technique of double GAN and convolutional autoencoder to the detection of abnormalities in chest X-rays.
2. The network is optimised using the error loss between the original and reconstructed samples, and it only requires normal samples for training.
3. The network model presented in this research has a better generalisation ability and a higher detection accuracy than the previous advanced anomaly detection model for chest X-ray abnormality detection.

To develop a more complete system for identifying abnormalities in chest X-rays, we can further optimise the network model structure in the future depending on the features of chest X-ray pictures. Furthermore, by lowering network model parameters and cutting down on network operating time, the upcoming research paths seek to increase the effectiveness of real-world medical applications.

Author Contributions: Conceptualization, Tan Yanli. And Azliza Mohd Ali.; methodology, Tan Yanli.; software, Tan Yanli.; validation, Tan Yanli., Azliza Mohd Ali. and Sharifalillah Nordin.; formal analysis, Tan Yanli.; investigation, Wang Jin.; resources, Li Guoqin.; data curation, Tan Yanli.; writing—original draft preparation, Tan Yanli.; writing—review and editing, Sharifalillah Nordin.; visualization, Tan Yanli.; supervision, Li Guoqin.; project administration, Tan Yanli.; funding acquisition, Tan Yanli. All authors have read and agreed to the published version of the manuscript”.

Funding: This research received no external funding.

Acknowledgments: The authors would like to express gratitude to the School of Computing Sciences, College of Computing, Informatics, and Mathematics for all the support given.

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

REFERENCES

- [1] Li P, Pei Y, Li J. A comprehensive survey on design and application of autoencoder in deep learning[J]. *Applied Soft Computing*, 2023, 138: 110176.
- [2] Zhao P, Ding Z, Li Y, et al. SGAD-GAN: Simultaneous Generation and Anomaly Detection for time-series sensor data with Generative Adversarial Networks[J]. *Mechanical Systems and Signal Processing*, 2024, 210: 111141.
- [3] Liu R, Liu W, Zheng Z, et al. Anomaly-GAN: A data augmentation method for train surface anomaly detection[J]. *Expert Systems with Applications*, 2023, 228: 120284.
- [4] Lian Y, Geng Y, Tian T. Anomaly detection method for multivariate time series data of oil and gas stations based on digital twin and mtad-gan[J]. *Applied Sciences*, 2023, 13(3): 1891.
- [5] Torabi H, Mirtaheeri S L, Greco S. Practical autoencoder based anomaly detection by using vector reconstruction error[J]. *Cybersecurity*, 2023, 6(1): 1.
- [6] Mattera G, Vozza M, Polden J, et al. Frequency informed convolutional autoencoder for in situ anomaly detection in wire arc additive manufacturing[J]. *Journal of Intelligent Manufacturing*, 2024: 1-16.
- [7] Fu H, Liang F, Liang J, et al. Asymmetric learned image compression with multi-scale residual block, importance scaling, and post-quantization filtering[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(8): 4309-4321.
- [8] Zhang Z, Wu H, Zhao H, et al. A novel deep learning model for medical image segmentation with convolutional neural network and transformer[J]. *Interdisciplinary Sciences: Computational Life Sciences*, 2023, 15(4): 663-677.
- [9] Lee J H, Kim J, Kwon S J, et al. Flexround: Learnable rounding based on element-wise division for post-training quantization[C]//*International Conference on Machine Learning*. PMLR, 2023: 18913-18939.
- [10] Eom K Y, Min B S. Development of deep learning network based low-quality image enhancement techniques for improving foreign object detection performance[J]. *Journal of Internet Computing and Services*, 2024, 25(1): 99-107.
- [11] Rachmadi M F, Byra M, Skibbe H. A new family of instance-level loss functions for improving instance-level segmentation and detection of white matter hyperintensities in routine clinical brain MRI[J]. *Computers in Biology and Medicine*, 2024, 174: 108414.
- [12] Chmielewska A, Dean D. The role of stiffness-matching in avoiding stress shielding-induced bone loss and stress concentration-induced skeletal reconstruction device failure[J]. *Acta Biomaterialia*, 2024, 173: 51-65.
- [13] Szymanowicz S, Rupprecht C, Vedaldi A. Splatter image: Ultra-fast single-view 3d reconstruction[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024: 10208-10217.
- [14] Cai D, Jia T, Wang H, et al. XTM-GAN: Generative Adversarial Networks for Tone Mapping of Dual-Energy X-ray Security Images[C]//*2024 IEEE 14th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 2024: 416-421.
- [15] Li X, Wu W, Yan W, et al. Randomly-Accessible Active Pixel Sensor with Logarithmic-Exponential Response Enabling Low Dose and Wide Dynamic Range Indirect-Conversion X-ray Imaging[J]. *IEEE Sensors Journal*, 2024.
- [16] Singh R, Sethi A, Saini K, et al. Attention-guided generator with dual discriminator GAN for real-time video anomaly detection[J]. *Engineering Applications of Artificial Intelligence*, 2024, 131: 107830.
- [17] Hassan E, Shams M Y, Hikal N A, et al. The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study[J]. *Multimedia Tools and Applications*, 2023, 82(11): 16591-16633.
- [18] Shah P S, Hughes E G, Sukhadia S S, et al. Validation and implementation of a somatic-only tumor exome for routine clinical application[J]. *The Journal of Molecular Diagnostics*, 2024, 26(9): 815-824.
- [19] Hui V, Litton E, Edibam C, et al. Using machine learning to predict bleeding after cardiac surgery[J]. *European Journal of Cardio-Thoracic Surgery*, 2023, 64(6): ezad297.
- [20] Bakurov I, Buzzelli M, Schettini R, et al. Full-reference image quality expression via genetic programming[J]. *IEEE Transactions on Image Processing*, 2023, 32: 1458-1473.
- [21] Andono P N, Setiadi D R I M. Quantization selection based on characteristic of cover image for PVD Steganography to optimize imperceptibility and capacity[J]. *Multimedia Tools and Applications*, 2023, 82(3): 3561-3580.
- [22] Pu B, Lan S, Wang W, et al. GanNeXt: A New Convolutional GAN for Anomaly Detection[C]//*International Conference on Artificial Neural Networks*. Cham: Springer Nature Switzerland, 2023: 39-49.
- [23] Liu J, Xie G, Wang J, et al. Deep industrial image anomaly detection: A survey[J]. *Machine Intelligence Research*, 2024, 21(1): 104-135.



© 2025 by Tan Yanli, Azliza Mohd Ali, Sharifalillah Nordin, Wang Jin, and Li Guoqin. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).