



Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com



An optimal defect recognition security-based terahertz low resolution image system using deep learning network

Samuel Akwasi Danso ^{a,b,*}, Shang Liping ^b, Deng Hu ^b, Samuel Afoakwa ^a, Eugene Louis Badzongoly ^a, Justice Odoom ^b, Owais Muhammad ^b, Muhammad Umer Mushtaq ^b, Abdul Qayoom ^b, Wenqing Zhou ^c

^a Ghana Communication Technology University, Ghana

^b Southwest University of Science and Technology-Mianyang -Sichuan, China

^c South China University of Technology, Guangzhou, China



ARTICLE INFO

Article history:

Received 22 June 2021

Revised 6 May 2023

Accepted 14 May 2023

Available online 29 May 2023

Keywords:

Defect recognition

Differential evolution search algorithm

Feature pyramid network

Generative adversarial network

Low resolution

Terahertz image

ABSTRACT

The physics of Terahertz (THz) technology is the electromagnetic (EM) spectrum band between the infrared and the microwave band with frequencies of about 0.1 to 30 THz. THz signals have gained adoption in medicine, telecommunications, security monitoring and imaging. THz imaging technology has the advantages of rapid imaging, strong penetration, and harmless to the human body hence widely used in a variety of security environments and has become an alternative technology for X-ray imaging. However, THz is characterized by low resolution of THz images of which noise is an integral factor constituting a defect. Clarity of THz image is therefore essential at various security checkpoints to avoid life's dangers and treats. In this paper, we propose an efficient and high-performance defect detection model based on RetinaNet to recognize defects from captured images. The strategy of transfer learning is introduced to improve detection performance accuracy, which enhances the average precision (AP) by 19.2%. Contrary to existing THz image detection techniques on image recognition pertaining to the whole region of the image, we adopt a different approach via differential evolution search algorithm for optimization given the small proportion of defect area which improves the AP by 9.9% comparing with the fine-tuned model. For the problem of the lack of defect data samples, image augmentation is adopted to enrich our training samples, which improves the AP by 9.5%. As for the problem of low precision and recall in detecting blurred images, we firstly manually generate clear-blurred image pairs to train a GAN. Then, the blurred images are deblurred using a trained generator. We get 5.5% AP improvement on the testset using our approach. Compared with existing works, the optimized model based on RetinaNet has better detection performance, subsequently, proving the practicability and effectiveness of the proposed method.

© 2023 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Computers and Artificial Intelligence, Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Terahertz within the electromagnetic spectrum band (frequency range from 0.1 to 10 THz) encompasses both the electronics and photonics fields can penetrate a broader range of dielectric materials, including cloth, paper, plastic, leather, and wood. It is

non-ionizing and has only minor side effects on humans [1]. It has a high rate of absorption in water and a high rate of radiation reflection in metals and is currently used in a variety of fields, including medicine, security checkpoints, postal parcels etc. Despite this, Terahertz images have poor resolutions due to their low power source, resulting in blurred and noised images constituting a defect. Therefore, it is necessary to propose an accurate and efficient intelligent defects recognition method.

Deep learning models have significant influence on X-ray and terahertz images. Deep learning was the first to shine in the picture field, breaking ground in image classification, detection target or localisation, semantic segmentation, and many other areas [2]. In the past decades, with the proposal of the histogram of oriented gradient (HOG), scale-invariant feature transform (SIFT), and local binary pattern[3-5] on image recognition as well as image com-

* Corresponding author at: Ghana Communication Technology University, Ghana.

E-mail addresses: skdanso1@gmail.com, sdanso@gctu.edu.gh (S.A. Danso).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.



Production and hosting by Elsevier

pression [6] has received tremendous attention. The abstract features that can adequately describe objects are often ignored when designing features manually [7], and the designed network usually needs to be trained separately to conduct multi-level positioning. There is also the case of correctness of the labels consequently affecting performance of models [8], the problem of class imbalance [9], and the constrain of complex and challenging data collection [10]. We address these challenges in our work via a myriad of approaches including data augmentation which we elucidate later. In recent years, the boom of deep learning based on a convolutional neural network (CNN) provides another idea for object recognition [11]. Compared with the manual feature design method mentioned above, the CNN does not require the extraction of image features manually hence avoids the tedious process of manual feature design, and realizes end-to-end defect detection. Nevertheless, the above research outcomes still have some limitations. For instance, [12] employs evolutional state construction algorithm and an optimized evolutional observation algorithm to detect anomalous changes in social networks. The capability to detect 3D Symmetry from Single-View RGB-D images using weak supervision has also been proposed [13] as well as global and local-contrast guides content-aware fusion for RGB-D regarding saliency prediction [14]. Although great strides have been made, none of these focuses on defect detection in THz images.

The motivation for this work stems from the problems relating to resolution, blurriness and noise of THz images. Spurred on to resolve such challenges, we introduce RetinaNet, a kind of one-stage detector, to recognize the defects of the terahertz low resolution images used in security checkpoint and other real-life environments. ResNet50, an efficient CNN, is firstly used as the backbone to extract features. Then a feature pyramid network (FPN) is used to collect semantic information on multiple scales produced by ResNet50. In the end, the defects of terahertz low resolution images are identified by classification subnetwork and box regression subnetwork. Notice that unlike extant works, in this work, we utilize transfer learning and GAN, which can be regarded as a combination of CNN and GAN. Again, given that anchor configuration is crucial for the anchor-based detectors, we optimize default aspect ratios via differential evolution search algorithm.

We organize the remaining sections of this work as follows. Section 2 introduces other THz detection methods and the structure of the model for defect recognition whereas the GAN is introduced in Section 3, including the structure of the generator and the loss of GAN. Section 4 presents the processing defect detection of terahertz low resolution images combining with CNN and GAN. Section 5 introduces our experiments and results, including optimization of anchor configuration, data augmentation and comparison of other detection frameworks. Section 6 introduces the improvement approach for detecting blurred images using GAN. Section 7 summarizes the above contents.

2. Image recognition

We begin this section by introduction diverse Terahertz image recognition strategies. Subsequently, we advance details pertaining to defect recognition.

2.1. Terahertz Image Recognition

Recently terahertz technology and deep learning have been a focus of research. The applied areas in the domain of the field object detection are detection of agricultural products [15–18], breast cancer and other medical conditions [19–21], hiding object[22–25]. Nevertheless, terahertz suffers low image resolution due to blurriness and noisy darken-spotted images which

stems from low energy power sources[26–28] consequently affecting detection accuracy and rate. It, therefore, stands to reason that any attempt to increase the detection rate and accuracy must first address the challenge of low resolution while at the same time revamping the deep learning model. Following, we concisely explain some works in this regard.

2.1.1. Terahertz image Recognition in Machine Learning and waveform

Out of all the ML techniques under test, it was observed that random forest statistical algorithm works well with the THz datasets in both the frequency and time domains. With such ML algorithm, a classifier can be created with less than 1% out-of-bag error for segmentation. Moreover, random forest is practically the best choice among the several popular ML techniques under test for waveform recognition of time-domain data in terms of classification accuracy and timing [29]. To be able to learn both low-level extrinsic and high-level intrinsic features together, Huang et. al. [30] advanced a dual-graph attention convolution network (DGACN). Xie L. et. al. [31] also proposed a convolutional Siamese point net (CSPN) for partial-to-partial point cloud registration which embodied feature extraction, matching matrix computation and singular value decomposition (SVD) for optical measurement of blade profile. Yin M. et. al. [32] also advanced an approach in the same domain leveraging deep feature interaction network (DFINet) where a feature interaction module to strengthen information association between two point clouds during feature extraction is used.

2.1.2. Terahertz image recognition in Agricultural products

Li B. et. al. (2020). [33] use quantitative methods for identification of common agricultural diseases. Their exploratory study involved data analysis of common fungal pathogens using identification modeling based on terahertz spectrum technology. Using polyethylene as the control, the terahertz time-domain spectra, and frequency-domain spectra of samples of the three pathogens were both measured. The absorption and refraction characteristics of these samples in the range of 0.1–2.0 THz were calculated and analyzed, and samples were then divided using the KS algorithm. Terahertz spectrum-image data blocks of the pathogen samples were preprocessed, and the dimensions of data were reduced using non-local mean filtering and the SPA algorithm, respectively. K-nearest neighbors (KNN), support vector machine (SVM), and BP neural network (BPNN), and other algorithms were used for analysis of terahertz images at characteristic frequencies, and for investigating the identification model. The model was quantitatively evaluated, and its imaging visualization was studied with results suggesting that there are significant differences among diseases in absorption and refraction in the terahertz band.

2.1.3. Terahertz images recognition for Automatic target low-count images

The physical properties of terahertz radiation and associated hardware, such images have low contrast, signal-to-noise ratio and low resolution i.e. all the characteristics of a low-count images. Therefore, standard recognition algorithms designed for conventional images work poorly or are not suitable at all for the problem considered. A differential approach based on clustering 2D point clouds was developed in accordance with a set of predefined patterns. As a result, there was reduction in the problem of target recognition to the problem of maximizing the image data likelihood with respect to the classes of model objects up to the size and position. The resulting recognition algorithm has a structure close to that of the well-known EM algorithm according to Antsiperov, V. E. et al. (2016) [34].

2.1.4. THz image recognition in medical fields

Terahertz (THz) imaging has the potential to detect breast tumors during breast-conserving surgery accurately. Over the past decade, many research groups have extensively studied THz imaging and spectroscopy techniques for identifying breast tumors. Recent development of THz imaging techniques for breast cancer detection have been put forward Wang L [19] in which the dielectric properties of breast tissues in the THz range, THz imaging and spectroscopy systems, THz radiation sources, and THz breast imaging were studied. In addition, numerous chemo metrics methods applied to improve THz image resolution and data collection processing are summarized. Similarly multi-dimensional mining algorithms in relation to Magnetic Resonance Imaging (MRI) radiogenomics for computer aided detection and diagnosis of breast tumours was proposed Yin, Xiao-Xia, et al. [20] to address a new problem in radio genomics mining: how to combine structural radiomics information with non-structural genomics information for improving the accuracy and efficacy of Neoadjuvant Chemotherapy (NAC). It required the automated extraction of parameters from non-structural breast radiomics data, and finding feature vectors with diagnostic value, which then are combined with genomics data. In order to address the problem of weakly labelled tumour images, a Generative Adversarial Networks (GAN) based deep learning strategy is proposed for the classification of tumour types; this has significant potential for providing accurate real-time identification of tumorous regions from MRI scans. In order to efficiently integrate in a deep learning framework different features from radiogenomics datasets at multiple spatio-temporal resolutions, pyramid structured and multi-scale densely connected U-Nets are proposed. A bidirectional gated recurrent unit (BiGRU) combined with an attention based deep learning approach is also proposed. More recently, Zhuang Y. et. al. [35] employed Weakly Supervised Similarity Evaluation Network (WSSENNet) to efficiently support similarity analysis of lung CT images (CTIs) thereby guaranteeing higher retrieval accuracy. A similar recent work [36] on large CTI Sequences in Mobile Telemedicine Networks also exist.

2.1.5. Terahertz image in postal parcel recognition and mailing

The work by Sleiman, Joyce el al[37] demonstrates the potential and capabilities of terahertz technology for parcels screening and inspection to detect threats such as weapons and explosives, without the need to open the parcel. They present terahertz time-domain spectroscopy and spectral imaging for explosives detection. Two types of explosives as well as their binary mixture is analyzed. Due to the complexity of extracting information when facing such mixtures of samples, three chemo metric tools are used: principal component analysis (PCA), partial least square analysis (PLS) and partial least squares-discriminant analysis (PLS-DA). The analyses are applied to terahertz spectral data and to spectral-images in order to describe a set of unknown data and identify similarities between samples by PCA, create a classification model and predict the belonging of unknown samples to each of the classes, by PLS-DA, and create a model able to quantify and predict the explosive concentrations in a pure state or in mixtures, by PLS. Subsequently the work focused on millimeter wave imaging for weapon detection in parcels in which three different imaging techniques are studied: passive imaging, continuous wave (CW) active imaging and frequency modulated continuous wave (FMCW) active imaging. The performances, the advantages and the limitations of each of the three techniques, for parcel inspection, are exhibited. Moreover, computed tomography is applied to each of the three techniques to visualize data in 3D and inspect parcels in volume. Thus, a special tomography algorithm is developed by taking in consideration the Gaussian propagation of the wave. Similarly, A Mail Inspection Based on Terahertz Time-Domain Spectroscopy

which is a spectral investigation of materials covered by visibly opaque objects. Terahertz waves are well suited to inspect the content of mail[38].

2.1.6. Terahertz image recognition for hidden or conceal objects

While malicious attacks on electronic devices (e-devices) have become commonplace, the use of e-devices themselves for malicious attacks has increased (e.g., explosives and eavesdropping). Modern e-devices (e.g., spy cameras, bugs or concealed weapons) can be sealed in parcels/boxes, hidden under clothing or disguised with cardboard to conceal their identities (named as hidden e-devices hereafter), which brings challenges in security screening. Inspection equipment (e.g., X-ray machines) is bulky and expensive. Moreover, screening reliability still rests on human performance, and the throughput in security screening of passengers and luggage is very limited. A low-cost and practical hidden e-device recognition technique was proposed Li, Zhengxiong, et al. [39] to enable efficient screenings for threats of hidden electronic devices in daily life. They investigated and modelled the characteristics of nonlinear effects, a special passive response of electronic devices under millimeter-wave (mmWave) sensing. Based on this theory and preliminary experiments, they designed and implemented, E-Eye, an end-to-end portable hidden electronics recognition system at frequency of 24 GHz mmWave probe, a smartphone-based e-device recognizer. To validate the E-Eye performance, they conducted experiments with 46 commodity electronic devices under 39 distinct categories.

2.1.7. Other techniques and application fields

Aside from the use of ML and deep learning techniques, other researchers have explored other techniques. Liu et. al. [40] employed square even-order derivative method coupled with numerical and experimental analysis to detect thermal defects in vehicle cables. To provide spatial location information and real-time tracking and adjusting the spatial position of surgical instruments, Liu et. al. [41] utilized 2D/3D multimode medical image registration to aid image-guided surgery using Gaussian Laplacian second-order differential operator. Zhou et. al. [42] also combined a CNN with transfer learning, data augmentation, and a Squeeze-and-Excitation Networks (SENet) for automated ore classification in mining and excavation.

The focus of all the existing THz image detection techniques is on image recognition pertaining to the whole region of the image. However, considering the small proportion of defect area, we adopt a different approach in this work by using differential evolution search algorithm for optimization. Following, we elucidate on our adopted approach.

2.2. Model for defect recognition

Traditional artificial neural networks often show excellent performance in simple classification and regression tasks, as shown in Fig. 1. According to their structure, they can be divided into the input layer, hidden layers and output layer, which play the role of receiving data, processing data, and producing results as output [43,44]. However, in the field of image recognition, the traditional artificial neural network is gradually being replaced by a convolutional neural network due to the excessive number of parameters and the inability to make the best of the position information between pixels. The convolutional neural network has been widely used in the field of image recognition relying on its unique design ideas, such as local connection, weight sharing, and down-sampling, etc. The conventional convolutional neural network has the following structures: convolutional layer, pooling layer, and fully-connected (FC) layer. An activation function is set in front of the pooling layer [7].

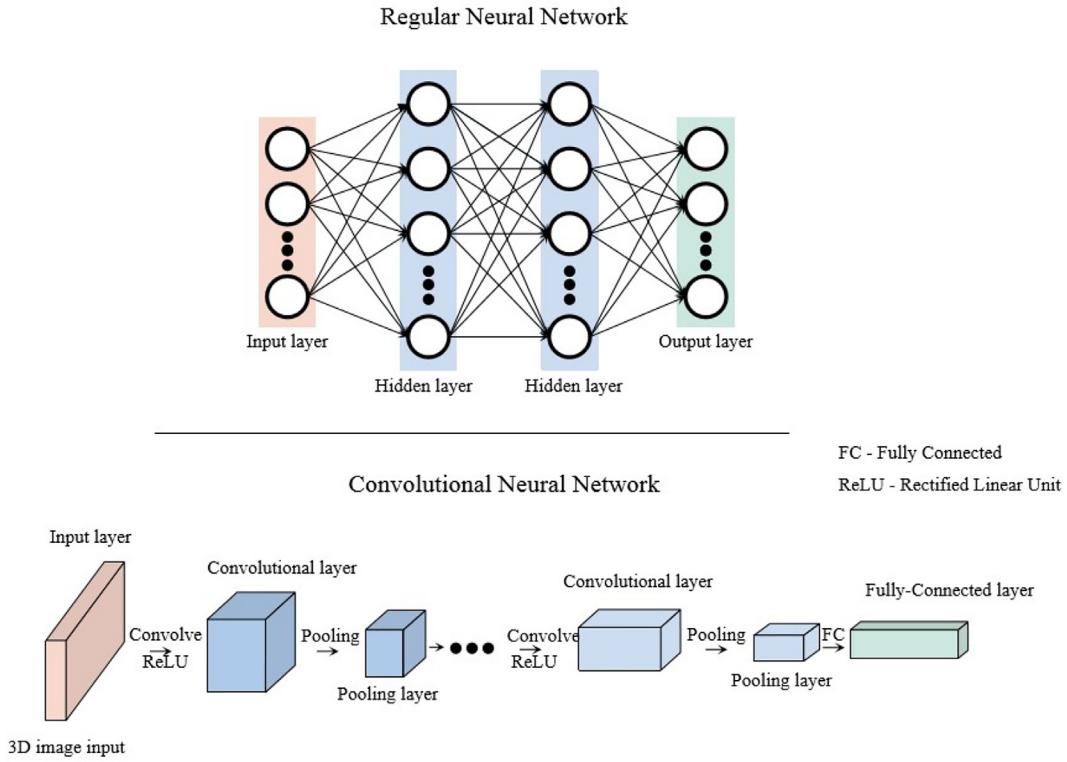


Fig. 1. Structure of traditional artificial neural network and convolutional neural network.

2.2.1. Structure of the defect detection model

Considering the high performance of detection, RetinaNet [45] is used as our basic framework to detect the defects of the terahertz low resolution images. The structure of the defect detection model is shown in Fig. 2. The main structure includes a backbone network and two subnetworks. Among them, the backbone network adopts a feature pyramid network (FPN) [46,47] to calculate the feature map of the whole image. The FPN constructs a rich, multi-scale feature pyramid through the top-down pathway and

lateral feature fusion. Considering that the residual network structure can well solve the problems of gradient disappearance and network degradation that are prone to occur in deep networks [25], the ResNet-50 is used as the backbone for feature extraction of FPN. The two subnetworks perform the box regression task and the classification task respectively at each output position of the backbone network. The classification subnet contains four 3×3 sizes of convolution kernels, equipped with 256 channels, followed by a rectified linear unit (ReLU) as well as a 3×3 convolu-

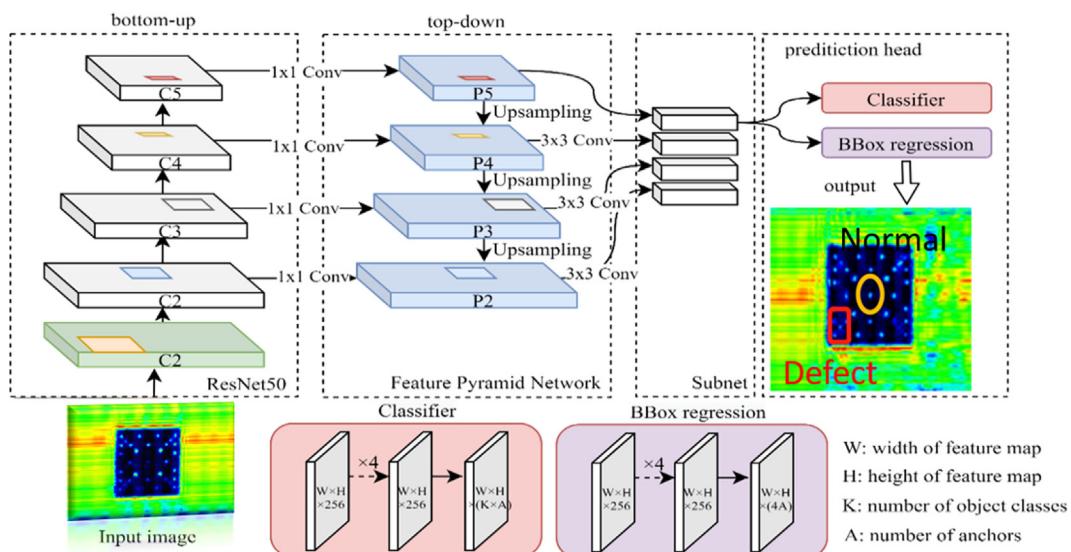


Fig. 2. Structure diagram of the defect detection model.

tional layer. Finally, the sigmoid activations are attached to the output layer. The box regression subnet is similar in structure to the classification subnet.

2.2.2. Loss function of classification and box regression

The object detector can be mainly classified into two types according to its structure and working principles. One type of object detector is based on a two-stage, proposal-driven mechanism, such as Faster-R-CNN [48] and Mask R-CNN [49]. Firstly, the region proposal networks (RPN) are used to generate a series of candidate object locations, on which the accurate classification and regression are carried out. The other type of object detector is one-stage. The typical regular one-stage detectors are SSD [50,51] and YOLO [29,30] [52,53]. Since time-consuming RPN are not adopted in this kind of detector structure, the inference speed of a one-stage detector is usually higher than that of a two-stage detector. Lin T. and others thought that the reason for the low precision of a one-stage detector is the class imbalance of the candidate locations. For each image to be detected, thousands of candidate locations will be generated. However, only a few of the numerous candidate locations contain detection objects, and most of the candidate locations are easily distinguishable negative samples. When using the regular cross entropy (CE) loss as the loss function for the task of classification, as shown in Eq. 1, the class imbalance will not only limit the model training rate but also lead to model degeneration.

$$CE(p, y) = CE(p_t) = -\log(p_t) \quad (1)$$

where p_t in the equation is defined as follows:

$$CE(p, y) = CE(p_t) = -\log(p_t) \quad (2)$$

where $p \in [0, 1]$. To reduce the negative effects brought by class imbalance, the focal loss is adopted as the loss function for the task of classification in this paper, whose expression is shown as follows.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (3)$$

where, α_t is the weighting factor greater than zero, and γ is the tunable focusing parameter. In Eq. 3, the larger the value of p_t , the lower the corresponding loss value. In other words, due to the presence of γ , the contribution of easily distinguishable negative samples to the loss value decreases, thus inhibiting the class imbalance of candidate regions. Since it is a defect detection task, in addition to determining what kind of defect, it is also necessary

to predict the defect location. Minimizing the bounding box regression loss can make the prediction result as close to the true value as possible. The loss used in the model is smooth L1 loss, which is shown as follows

$$Smooth_{L1} = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (4)$$

Considering there are four corners in a box, the entire loss to be minimized is expressed as:

$$L_{loc}(t^u, v) = \sum_{i \in (x,y,h,w)} Smooth_{L1}(t_i^u - v_i) \quad (5)$$

Where the (x,y) is the coordinate in the lower-left position, h and w represent the width and height, respectively; $v = (v_x, v_y, v_h, v_w)$ represents the coordinate of ground truth box; $t^u = (t_x^u, t_y^u, t_h^u, t_w^u)$ represents the coordinate of the predicted box.

3. Generative adversarial networks for image deblurring

The generative adversarial networks (GAN) was first presented by Goodfellow et al. [54], which used two neural networks (generator and discriminator) to pit one against the other (thus the adversarial) in order to generate new and synthetic data [55]. The structure of the GAN is shown in Fig. 3. The generator in the figure receives blurred images to generate new clear image. The discriminator is used to take the real clear samples and the generated samples as inputs and tries to distinguish them correctly.

3.1. The structure of generator and discriminator

The generator built in this paper is structurally consistent with the generator proposed by Kupyn o. et al. [56], as shown in Fig. 4. We use the clear-blurred image pair as the input of the GAN model. The generator consists of convolution blocks, residual blocks and transposed convolution blocks, whose numbers are set at 3, 9 and 2, respectively. Each convolution block contains a convolutional layer, instance normalization layer and activation layer. There are two sizes of convolution kernels to perform convolution computation, namely 7×7 and 3×3 . In order to accelerate the training rate of networks, the convolution layer is followed by instance normalization layer [57]. Also, the ReLU is chosen for activation function [58]. Considering that the shallow network cannot

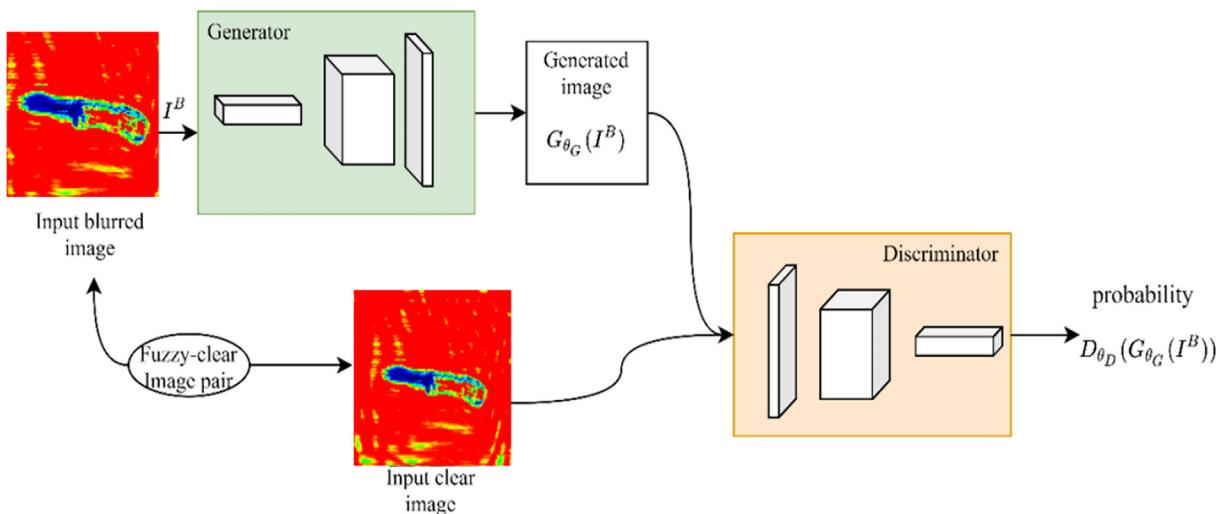


Fig. 3. Diagram of GAN for deblurring.

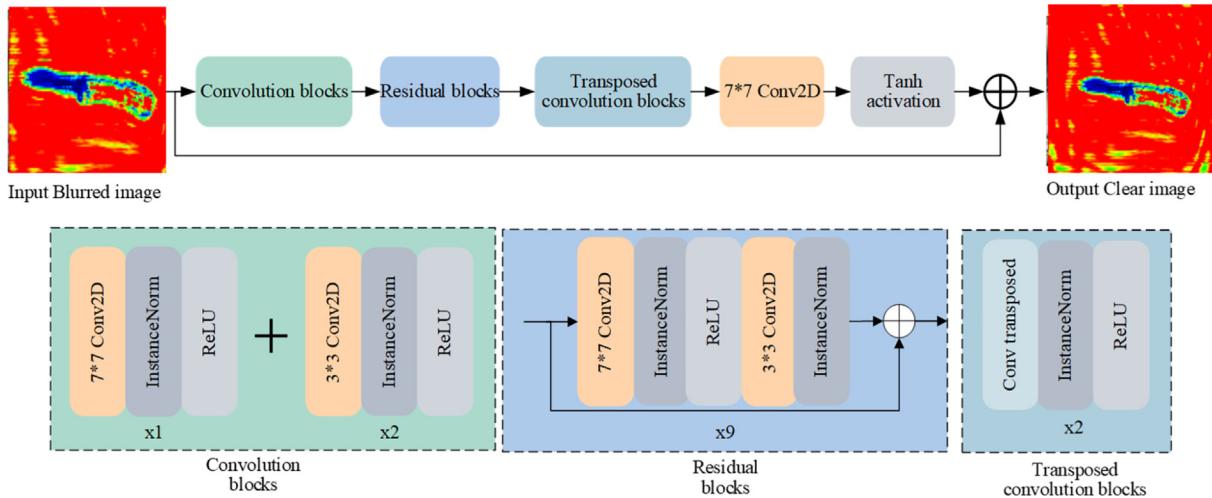


Fig. 4. Structure of Generator.

improve the performance of network recognition, and the stacking of conventional networks is prone to gradient disappearance the feature extraction is conducted by using the residual blocks. Each residual block contains two convolution layers, two instance normalization layers, and one activation layer. Finally, in order to obtain the reconstructed clear images, the transposed convolution blocks are adopted at the network terminal. In addition, in the output layer of the network, the global skip connection is adopted [56] because such formulation makes training faster and the resulting model generalizes better. The architecture of discriminator is identical to PatchGAN [59]. The LeakyReLU [60] is used as the function of the activation layer in all the remaining layers except the last convolution layer, and the corresponding parameter α is equal to 0.2 [56].

3.2. Loss Function of Generative Adversarial Networks

As the objective optimization function of GAN, the loss function plays a crucial role in the performance of the network. According to the architecture of the GAN network, the loss function is expressed as the sum of adversarial loss and content loss, as shown in Eq. 6.

$$L = L_{GAN} + \lambda L_X \quad (6)$$

Where, L_{GAN} stands for adversarial loss and, λL_X represents a content loss. According to the experimental results from Kupyn o. et al., the parameter λ is equal to 100. As for the selection of adversarial loss function, in order to generate high-quality images, the WGAN-GP is used as the critic function, which is shown to be robust to the choice of generator architecture [61]. The equation is shown as follows:

$$L_{GAN} = -\sum_{n=1}^N D_{\theta_D}(G_{\theta_G}(I^B)) \quad (7)$$

In this case, I^B is the input blurred image, and $D(\theta_D)(G(\theta_G)(I^B))$ represents the probability that the reconstructed images ($G(\theta_G)(I^B)$) output by the generated network are predicted as the actual images. The content loss usually takes the Mean-square Error (MSE) of the original pixel as the objective function of the optimization. Although the generated images have a high peak signal-to-noise ratio, it tends to lose high-frequency details, which makes the image tend to be smooth, even causing blurry artifacts due to the pixel-wise average of possible solutions in the pixel space [62]. In response to this problem, using perceptual loss proposed by Ledig C. and others to replace the mean-square error would be a good

choice. The perceptual loss is based on the difference between the generated image feature map and the target image feature map, hence the generated images and target images are more similar on semantic and style, which is defined as follows.

$$L_X = \frac{1}{W_{ij}H_{ij}} \sum_{x=1}^{W_{ij}} \sum_{y=1}^{H_{ij}} (\phi_{ij}(I^S)_{x,y} - \phi_{ij}(G_{\theta_G}(I^B))_{x,y})^2 \quad (8)$$

where $\phi_{i,j}$ is the feature map obtained by the j -th convolution (after activation) before the i -th maxpooling layer within the VGG19 network, pretrained on ImageNet. W and H represent the width and height of the feature map, respectively.

4. CNN-GAN combined model for defect recognition

Owing to the fact that objects in blurred image are difficult to detect alongside limited data samples, this paper proposes a method for detecting common defects in terahertz low resolution images based on transfer learning and GAN, which can be regarded as a combination of CNN and GAN. To begin with, considering that it is difficult to obtain the blurred-clear image pairs used for GAN training in practice, we generate blurred-clear image pairs based on the clear images by using Markov process [63,64] and subpixel interpolation [65–67]; Thereafter, we use the blurred-clear image pairs to train the generator and discriminator of GAN alternately until the specified epoch are reached. After finishing the training process of GAN. Reconstruction of the blurred images are carried out by the trained generator. Finally, the constructed clear images are used to fine-tune the backbone network, and the defected terahertz low resolution images are identified by the trained defect detection model. Fig. 5 shows the detailed implementation of the CNN-GAN combined detection model.

5. Experiments and results discussion

We trained and tested the model by using PyTorch, a kind of popular deep learning framework, with two RTX 2080 graphic cards on the server. The detailed hardware, software configuration and terahertz image acquisition device are listed in Appendix A. Other experimental results are included in Appendix A.

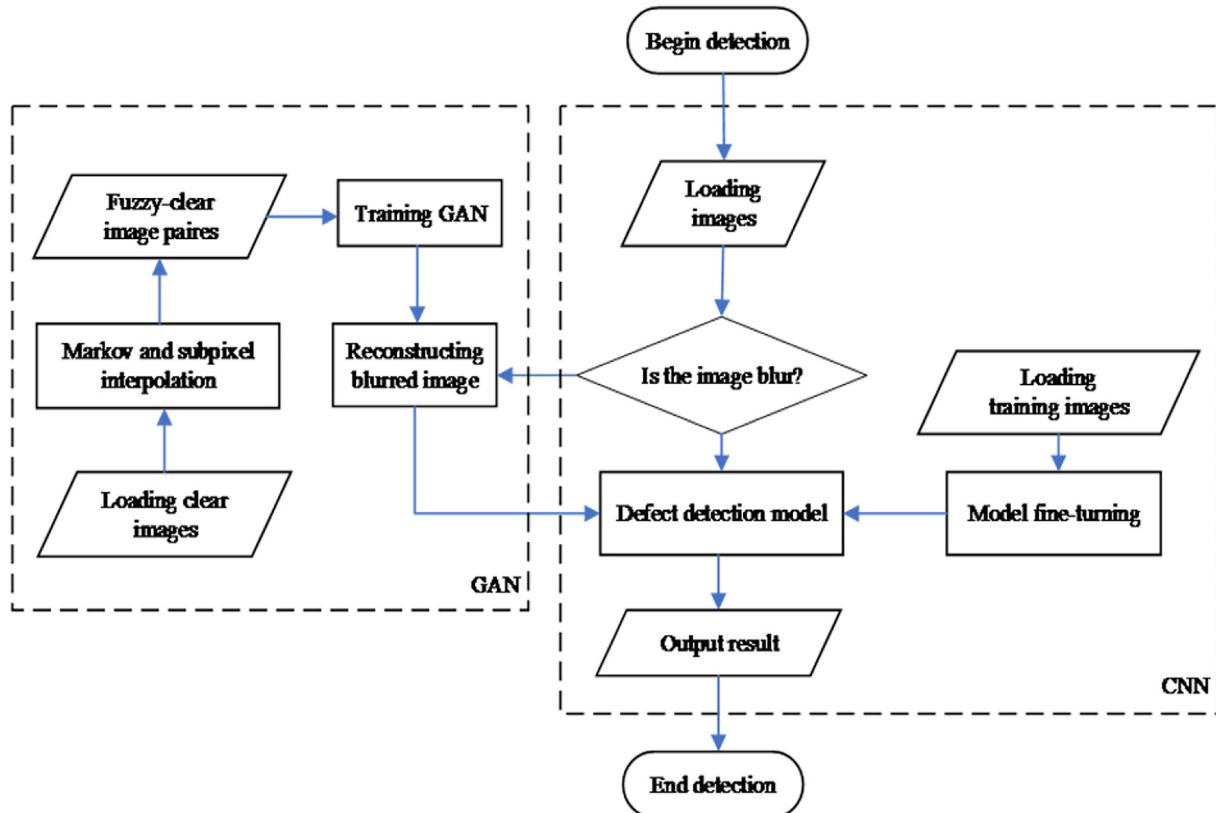


Fig. 5. Terahertz low resolution images defect detection framework based on CNN and GAN.

5.1. Metrics for Detection

In this paper, we adopted the detection metrics introduced in [68], which includes average precision (AP) and average recall (AR) over multiple Intersection over Union (IoU) values. The IoU can be calculated by:

$$\text{IoU}(Box_{pred}, Box_{gt}) = \frac{Box_{pred} \cap Box_{gt}}{Box_{pred} \cup Box_{gt}} \quad (9)$$

The calculations of precision and recall are shown in Fig. 6. Traditionally, AP is calculated in a single category, and mean average precision (mAP) is averaged over all categories (likewise AR and mAR).

Table 1
Dataset information for training and testing.

Metric	Meaning
AP	Average precision at IoU = 0.5:0.05:0.95
AP ^{0.5}	Average precision at IoU = 0.5 (PASCAL VOC metric)
AP ^{0.75}	Average precision at IoU = 0.75 (strict metric)
AP ¹	Average recall given 1 detection per image
AP ¹⁰	Average recall given 10 detection per image
AP ¹⁰⁰	Average recall given 100 detection per image

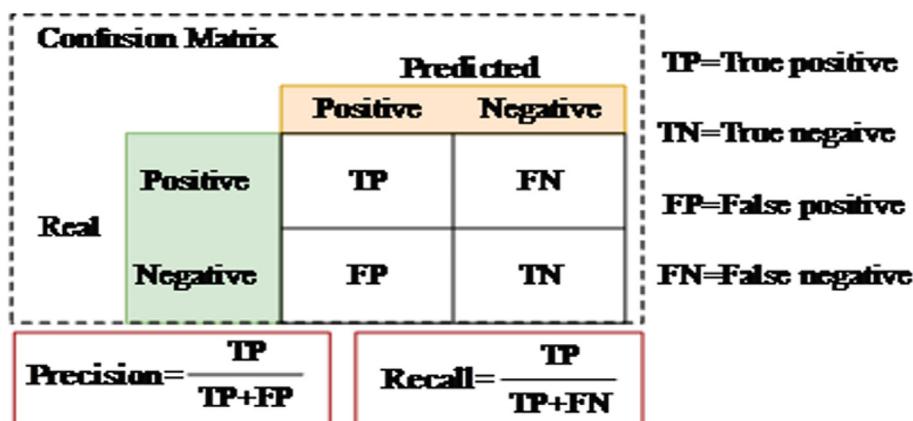


Fig. 6. Calculation of precision and recall.

Unless otherwise specified, AP and AR used in this paper are mAP and mAR. The detection metrics are listed in [Table 1](#).

AP (averaged across all 10 IoU thresholds and all categories) should be considered the most important metric when considering model performance in our research. For metric AR, the larger the value, the lesser the false negative rate, which is important for defect inspection along the overhead transmission line.

5.2. Datasets and Training configuration

The image dataset used in the experiment is collected by the terahertz linear scan machine. For this paper, we selected 1209 images in total for training and testing. These images were classified into two categories according to the state of the terahertz low resolution images: defected and normal. By using the labelling tool, the dataset was created with COCO [\[68\]](#) format. Finally we randomly divided the data samples into the training set and test set according to a ratio of nearly 5:1, as shown in [Table 2](#), and the statistical information of the total original dataset is shown in [Fig. 7](#). According to the statistical results, the average size of the labeled bounding box is 78px by 70px, accounting for 1.5% of the entire image area. In addition, the minimum size of box is 26px* 27px, whereas the maximum box is 206px*198px.

To train the defect detection model introduced in previous Section 2, it is essential to setup the training configurations properly. In our research, we use the configurations listed in [Table 3](#).

For all the models conducted in this paper, the configuration of training remain the same to guarantee the validity of the research.

5.3. Defect detection of terahertz low resolution images based on transfer learning

The convolution neural network can extract the latent features in images efficiently. After being trained by using an image dataset (like ImageNet), the weights of convolutional network obtain common features representing semantic information. The idea of transfer learning is to use the learned network weights to fine-tune the new detection network. It converges quicker comparing with the method of training from scratch. The model results of AP and AR during training are shown in [Fig. 8](#). The evaluation results on the test dataset are listed in [Table 4](#).

[Fig. 7](#) reveals that there has been a remarkable increase in the performance of the fine-tuned model. For the evaluation on test dataset, the metric AP also improved by 19.2% and the metric AR100 improved by 15.7%.

5.4. Optimize anchor configuration using differential evolution search algorithm

The anchor configuration is crucial for the anchor-based detector. In the original model, the default aspect ratios and scales of the anchor box were (1:2, 1:1, 2:1) and (20/3, 21/3, 22/3), respectively. In such configuration, the model achieved a good result on COCO dataset. In this experiment, we optimized these configurations using differential evolution search algorithm for our own terahertz low resolution images defect dataset. The objective is to maximize the overlap between the lesion bounding-box and the best anchor on the validation dataset. We fix one ratio as 1:1, and define other

Table 2
Dataset information for training and testing.

Target Label	Training	Testing	Total
Normal	660	129	789
Defected	545	111	656
Total	1205	240	1445

ratios as reciprocal pairs. The algorithm for searching the optimized anchor configurations is listed in Algorithm 1.

Algorithm 1: Optimize anchor ratios and scales.

Input: Labeled bounding box
(x_1, y_1, x_2, y_2)
Output: Optimized ratios, scales
Data: **Require** N_{ratio} : number of ratio to be optimized, set it to be 3;
Imgshape: Image shape, 600px*600px; Miter: maximum iteration, default is 1000

```

1  $L = \text{size}(\text{labeledboundingbox})$ 
2 Initialize ratio bound to [1,4], scale bound to [0.4,2]
3 iteration=0, tolerance=1e - 5
4 for  $i = 1$  to  $L$  do
5   append(entries, ( $x_1, y_1, x_2, y_2$ )) # ( $x_1, y_1$ ), ( $x_2, y_2$ ) are coordinate of bounding box
6 end
7 function(ratios, scales, Nratio, Imgshape)
8 anchor=gen_anchor(ratios, scales,  $N_{ratio}$ , Imgshape)

9  $OL = \text{cal\_overlap}(\text{entry}, \text{anchor})$ 
10  $\text{result} = -(1 - OL)2 * \log(OL)$  # maximize it
11 return result
12 end function
13 for inter <  $M_{inter}$  do
14   differential_evolution(function(*args),
15   ratio_bound, scale_bound)
15 end
16 return ratios, scales

```

Finally we obtained the optimized anchor ratios values to be (0.911, 1.0, 1.234) and the optimized scales to be (1.275, 1.606, 2.0). Taking the find-turned model as our baseline model in this experiment, we then obtained the model results using optimized anchor configurations, which is shown in [Fig. 9](#) and [Table 5](#).

For the training process, there is a change of performance at epoch 36. After the model trained, the performance of AP and AR100 has been improved by 9.9% and 8.3%, respectively, compared with the baseline model.

5.5. Data Augmentation

Terahertz is a novel technology henceforth the image database is scanty. Therefore the number of terahertz image database sets that individually can collect would not be much as the regular database set for object detection model, such as PASCAL VOC, COCO, etc., hence will reduce the robustness and performance of the model. The images used in this paper are obtained using Tera-sense corporation terahertz linear scan machine. In this experiment, we augmented the training by using combination of translation, rotation, perspective transformation, affine transformation, mix-up, dropout of region, blurring, sharpen, contrast change, saturation change, cropping and etc. After augmenting, we had 1566 instances of normal and 1307 instances of defected.

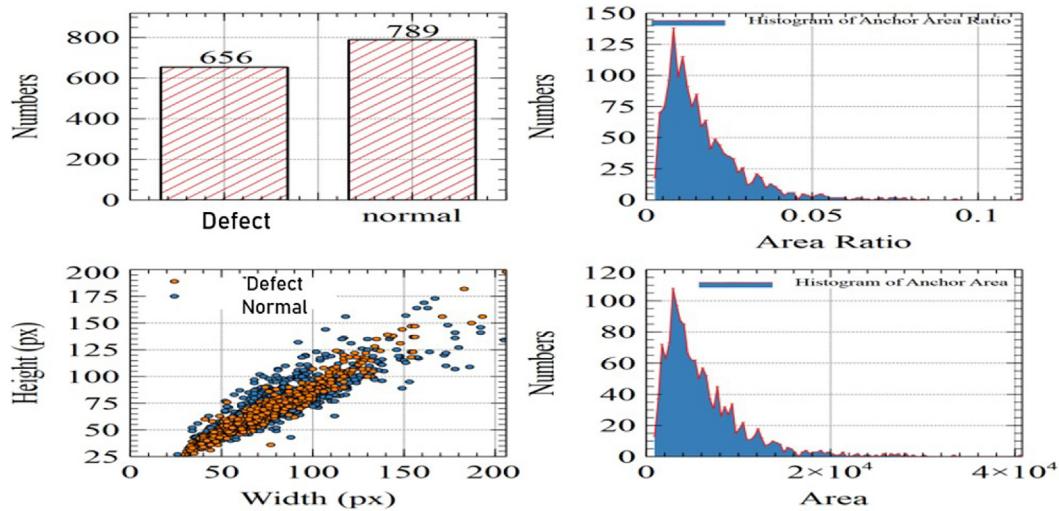


Fig. 7. Information on the original dataset.

Table 3
Configuration of for model training.

Parameter	Value
Training	epochs 50
Optimizer	Adam
Learning rate	1e-4
Batch size	8

Fig. 10 shows the training results using augmented training dataset. The evaluation on test dataset is listed in Table 6.

After augmentation, it is evident that the detection model has learned more features about the terahertz low resolution images, which improves the performance of model. Comparing to the results of baseline model, the average precision has been improved by 9.5%, and the average recall (given 100 detections per image) has been improved by 7.8%.

5.6. Comparison of other detection models

In addition to optimizing the model proposed in this paper, we also perform a comparison with other classic object detection algorithms, including Faster RCNN, SSD and YOLOv3. Note that YOLOv3

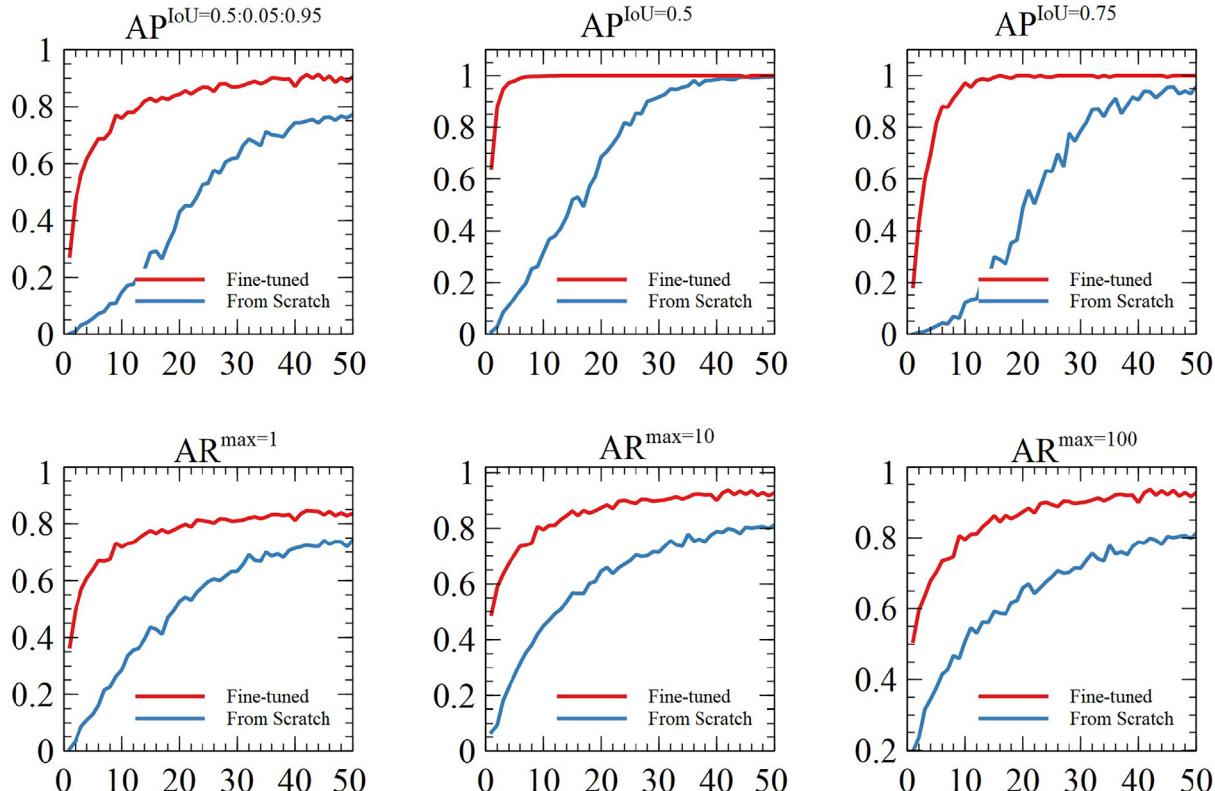
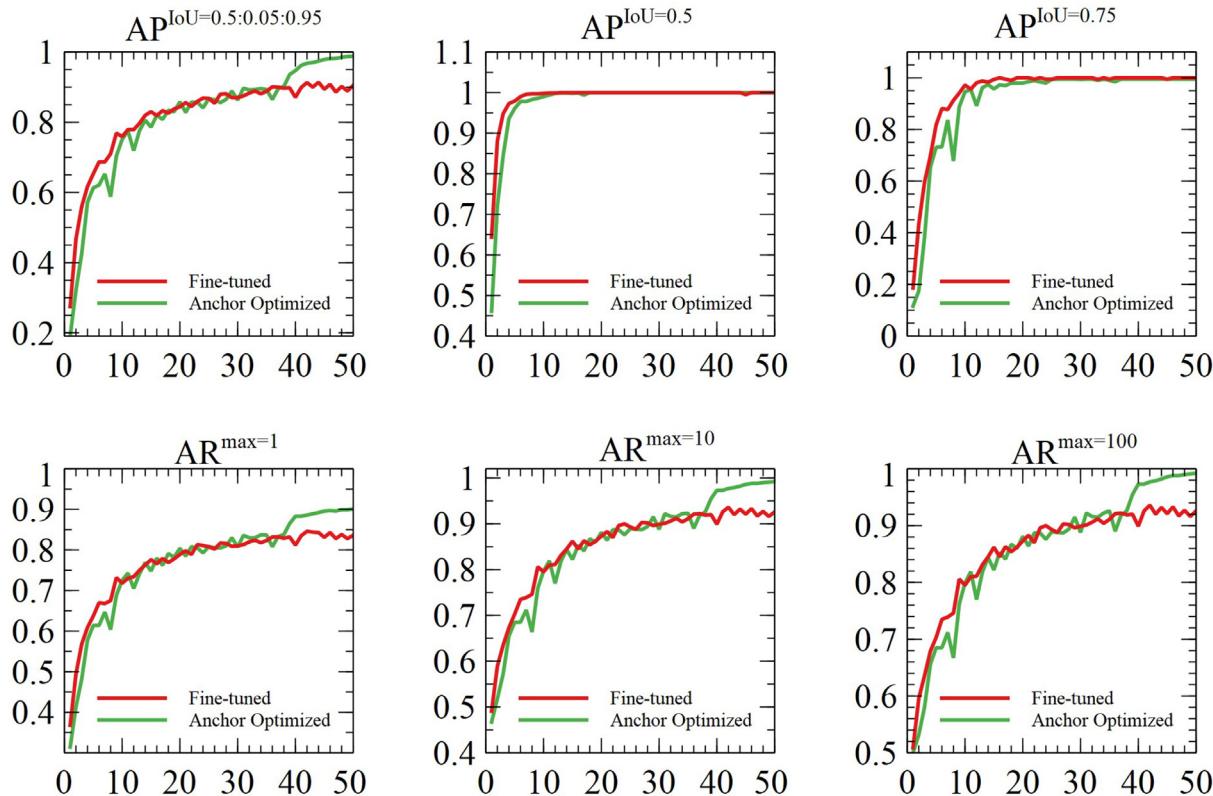


Fig. 8. Training results of model with and without fine-tuning.

Table 4

Evaluation results on test dataset.

Model	AP	$AP^{0.5}$	$AP^{0.75}$	AR^1	AR^{10}	AR^{100}
From Scratch	0.688	0.892	0.845	0.685	0.739	0.739
Fine-tuned	0.82	0.964	0.915	0.788	0.855	0.855
% ↑	19.2	8.1	8.3	15.0	15.7	15.7

**Fig. 9.** Training results of model with optimized anchor configuration.**Table 5**

Evaluation results on test dataset with optimized anchor configuration.

Model	AP	$AP^{0.5}$	$AP^{0.75}$	AR^1	AR^{10}	AR^{100}
Baseline	0.82	0.964	0.915	0.788	0.855	0.855
Anchor Optimized	0.901	0.977	0.939	0.845	0.926	0.926
% ↑	9.9	1.3	2.6	7.2	8.3	8.3

is chosen other than latest models like YOLOv4 and YOLOv5 in that YOLOv3 is still one of the most widely used detectors in industry due to its broad compatibility. Besides, there exist controversies (example short released time after yolov4, invalid benchmark dataset etc.) surrounding the capabilities of YOLOv5 [69].

The training parameters are consistent with Table 3, and the pre-trained weights are used for fine-tuning. The evaluation results of each model in the test set are listed in Table 7. Fig. 11 shows the visualization results, where the size of the circle represents the time spent in training the model.

Combined with the content of Table 7 and Fig. 10, it can be seen that the performance of model SSD500 is closer to the RetinaNet. However, the training time of the model SSD500 is 3.7 times than RetinaNet. The performance of the SSD300 is the worst because the model input image was scaled from 600px*600px to 300px*300px, which cause the loss of feature information. The model results of Faster RCNN and YOLOv3 are relatively close, but as Faster RCNN

belongs to the two-stage detector, its training time is longer, which is 2.4 times that of YOLOv3 model. This experiment also verified that the RetinaNet model has certain advantages in the terahertz low resolution images detection task, mainly because the focal loss for classification alleviates the problem of imbalance between positive and negative samples during the training process.

5.7. Deblur blurred Images using Generative Adversarial Network

In order to further improve the performance of our terahertz low resolution images defect detection model, we analyzed the output images produced by the fine-tuned model. We found that, for most of the higher resolution images, as shown in Fig. 12A, the detection model was able to not only accurately find the position of the terahertz low resolution images, but also identify the category (defected or normal) they belong. However, for some blurred images, the test show unsatisfactory results as shown in

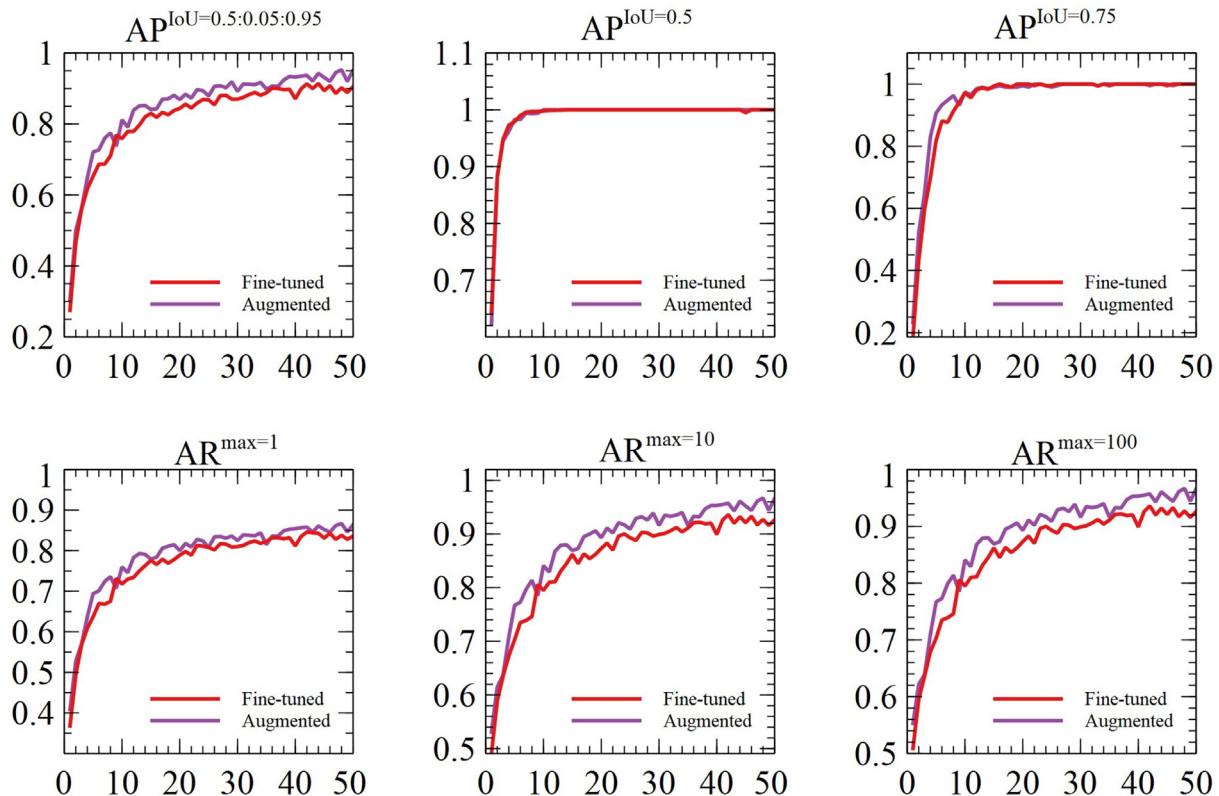


Fig. 10. Training results of model with augmented data.

Table 6

Evaluation results on test dataset with augmented data.

Model	AP	$AP^{0.5}$	$AP^{0.75}$	AR^1	AR^{10}	AR^{100}
Baseline	0.82	0.964	0.915	0.788	0.855	0.855
Data Augmented	0.898	0.995	0.967	0.842	0.922	0.922
% ↑	9.5	3.2	5.7	6.8	7.8	7.8

Table 7

Evaluation results on test with different defect detection models.

Model	AP	AR^{100}	Training Time
Fine-tuned	0.82	0.855	2.6 h
From scratch	0.688	0.739	2.6 h
Faster RCNN	0.727	0.786	7.6 h
SSD300	0.714	0.75	7.2 h
SSD500	0.812	0.846	9.6 h
YOLOv3	0.75	0.773	3.1 h

Fig. 12B. It can be seen from the results that the sharpness of this kind of image resolution is much lower than that of the image itself as shown in Fig. 12A.

To solve the problem of low precision and recall for detecting blurred images, we used generative adversarial network to reconstruct blurred images. Before doing this, we need to prepare clear-blurred image pairs for training GAN. Considering that the terahertz power source is the main reason for the low sharpness of such images, we used the method introduced by Kupyn o.e.t al. [56] to generate synthetic clear-blurred image pairs. Firstly, the trajectory vector was generated through Markov process. Then, the sub-pixel interpolation was applied to the trajectory vector to generate the blur kernels. Finally, the different types of blur kernels were applied to clear images to generate corresponding

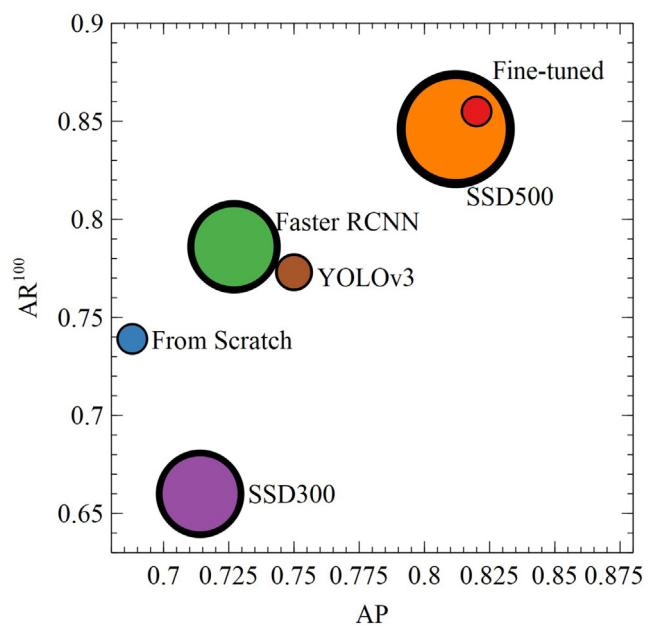


Fig. 11. Training results of the model with different defect detection models.

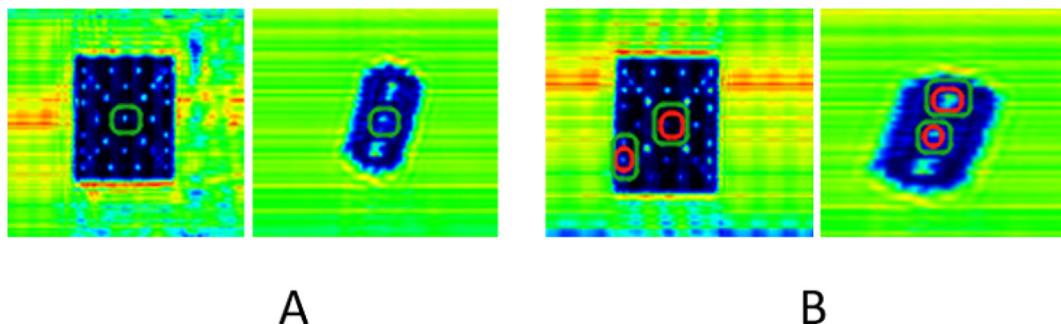


Fig. 12. A. Clear images that are correctly identified. B. Misidentified terahertz low resolution images in blurred image. The red box is the ground truth, and the green box is predicted by fine-tuned model.

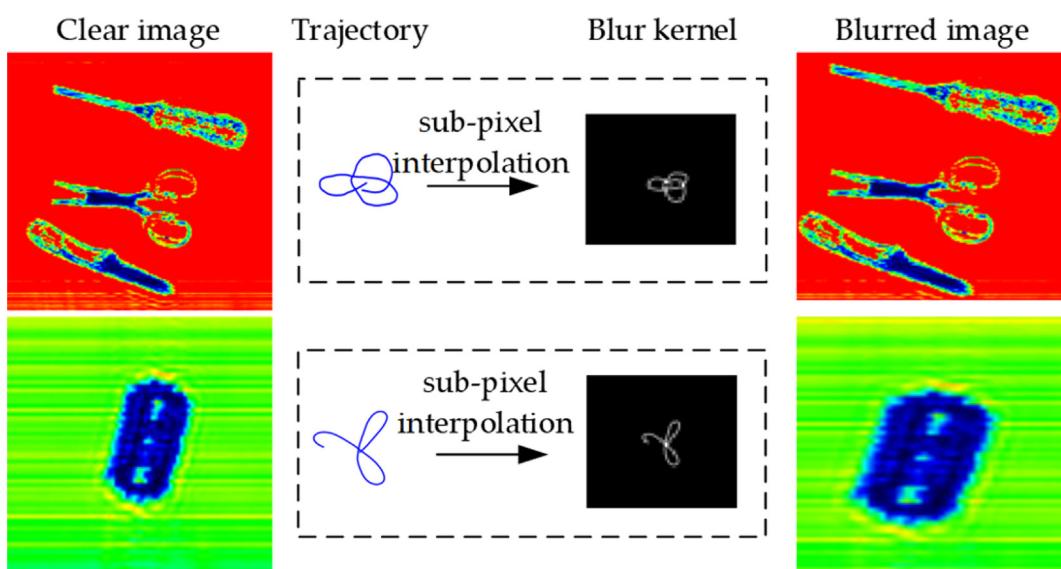


Fig. 13. Examples of generated clear-blurred image pair.

Table 8
Training configuration of GAN.

Parameter	Value
Training epochs	300
Optimizer	Adam
Learning rate	$[1e-4, 1e-5]$
Batch size	1

blurred images. **Fig. 13** presents the process of generation of clear-blurred image pairs.

We selected 590 clear images from the images collected by terasene corporation terahertz linear scan machine, and generated corresponding blurred images according to the above method. We then trained the GAN network model built in the original computer environment as per the parameters stipulated in **Table 8**. During the GAN training process, we performed 5 gradient descent steps on, then one step on, and used Adam as an optimizer to

Table 9
Evaluation results on deblurred testset with fine-tuned model.

Testset	AP	$AP^{0.5}$	$AP^{0.75}$	AR^1	AR^{10}	AR^{100}
Original	0.82	0.964	0.915	0.788	0.855	0.855
Deblurred	0.865	0.97	0.924	0.812	0.886	0.886
% ↑	5.5	0.6	0.9	3.0	3.6	3.6

Table 10
Evaluation results on different testsets with fine-tuned model.

Testset	$AP_{defected}$	AP_{normal}	AP	$AR_{defected}$	AR_{normal}	AR
Blurred	0.279	0.321	0.3	0.319	0.418	0.353
GAN reconstructed	0.358	0.426	0.392	0.387	0.495	0.457
% ↑	28.3	32.7	30.6	21.3	18.4	29.4

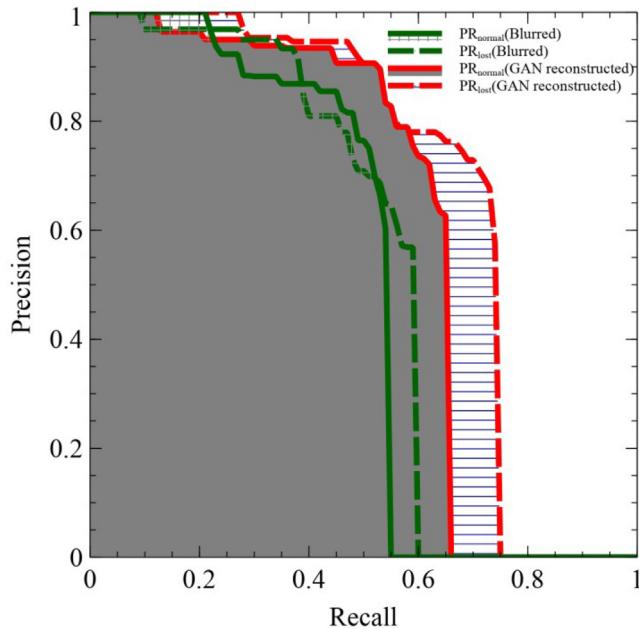


Fig. 14. Precision-Recall curve of normal category and defected category.

update the weight value. The training configuration is listed below. Note that the learning rate is set to $1e-4$ in the first 150 rounds, then we change it to $1e-5$ over the next 150 epochs.

When the training process of GAN is done, we reconstructed the blurred images in the original testset using the trained generator. Some results are shown in Table 9.

To verify the effectiveness of the method for reconstructing blurred images using GAN, the reconstructed dataset is evaluated by the fine-tuned model obtained before. The evaluation results are listed in Table 9.

Compared with the detection results in Fig. 13 and Table 9, after the original blurred image of the detection error is reconstructed by the generator, the detection model can give a better detection result. Table 9 shows that the metric AP has been improved by 5.5%, and the metric AR100 has been improved by 3.6% with the deblurred testset. To further verify the effectiveness of the method of deblurring fuzzy images, we generated corresponding blurred images using original testset (201 images). Then we reconstructed all the blurred images using trained generator. Finally we evaluated these two testsets to see their difference between the results as shown in the experimental results; Table 10 and Fig. 14.

It can be found that the average precision has been improved in both the normal category and the defected category in Fig. 14. The results from Table 10 also revealed that there been an improvement when reconstructed images are used to detect defects of terahertz low resolution images, because the average precision has been improved by 30.6%, and the average recall (given 100 detections per image) has been improved by 29.4%.

Table 11
Evaluation results on different testsets with fine-tuned model.

Hardware/Software	Parameters
Operating System	Ubuntu18.04 LTS 64bit (Linux 4.15)
Central Processing Unit	Intel(R) Core (TM) i7-7800X CPU @ 3.50 GHz
Graphical Processing Unit	NVIDIA RTX 2080(8G)
RAM	DDR4 32G
CUDA	CUDA 10.1
cuDNN	cuDNN 7.6.1
Deep Learning Framework	PyTorch 1.4

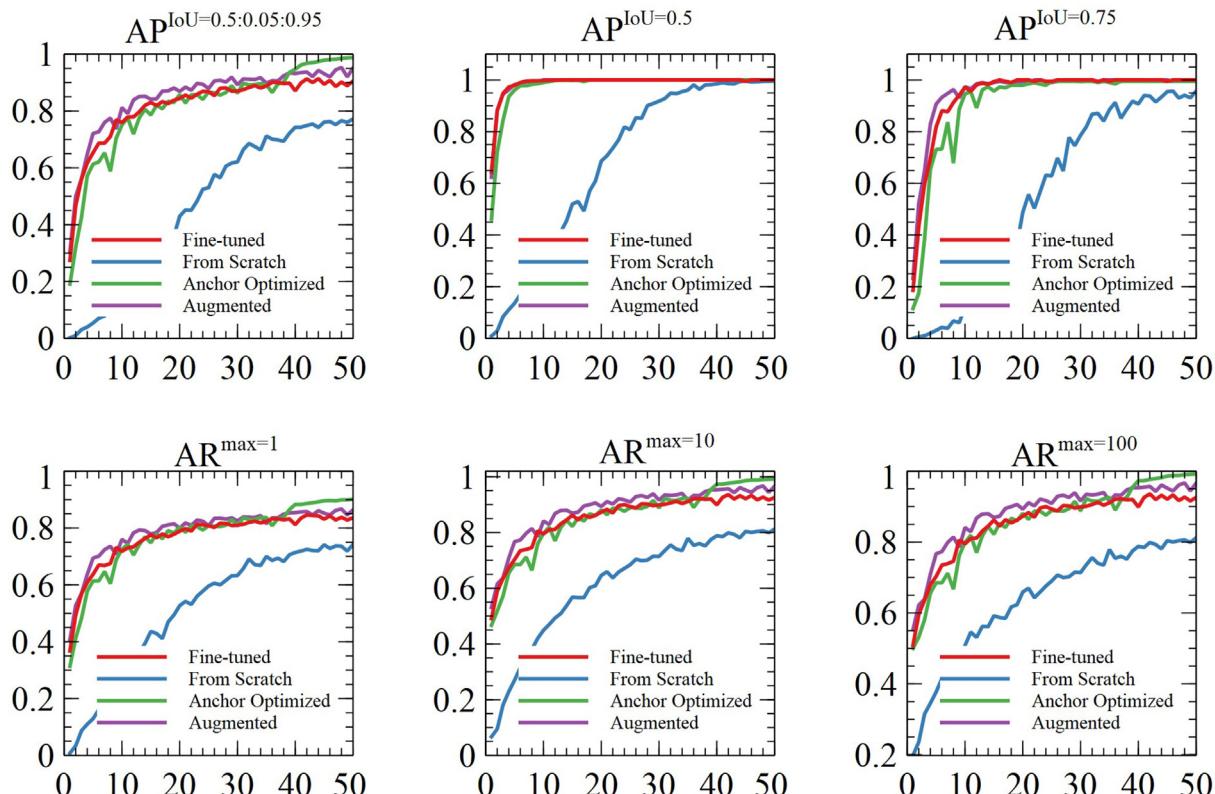


Fig. 15. Training results of four proposed models.

5.8. Discussion

Based on the experiment and results presented, it is obvious that after anchor optimization, the precision and recall of the proposed detection model significantly improves (see Table 5 and Fig. 9). The AP improvement witnessed sufficiently vouches for the effectiveness of the proposed anchor optimization algorithm. We therefore envisage the usage of the anchor optimization algorithm in similar image recognition tasks.

Also, examining the various augmentation techniques adopted, it is evident that the detection model learned more features inherent in the dataset avoiding any challenges like class imbalance consequently improving the performance of the model. For instance, comparing to the results of baseline model, the average precision improved by 9.5%. Moreover, from the results in Table 7 and Fig. 11, the proposed detection model outperforms state-of-the-art detection models including SSD300, SSD500, Faster RCNN and YOLOv3.

It is also noteworthy that the use of GAN by reconstructing blurred images can resolve the problem of low precision and recall pertaining to blurred images. However, it requires Markov process and sub-pixel interpolation. Our reliance on GAN as well as the

optimized anchor technique among others ultimately led to a better detection network for THz images.

6. Conclusions

In this paper, we proposed an efficient detection model for recognizing defects on terahertz low resolution images with improved methods. Transfer learning is used to fine turn the RetinaNet, which achieves a good performance improvement in average precision and average recall. Considering the small proportion of defect area, differential evolution search algorithm is used to optimize the anchor configuration for small defect detection. For the problem of lack of image data on terahertz low resolution images, we augment them manually by image processing, including translation, rotation, perspective transform, affine transform, and image mix-up. As for the problem of low precision and recall in blurred images, we reconstructed these images using the trained generator, which increases the performance of defect detection. In the future, we will focus on deploying our model on low-cost embedded devices, like Jetson Nano, to implement practical applications. **The Author's contributions, as follows:** S.A. gave the idea, did the experiments, interpreted the results and wrote the paper, S.L and D.H setup funding for THz scanning machine, S.A and E.L.B did result validation and review, J.O drew some of the figures and edited the paper, O.M, U.M and A.Q typed the manuscript, W.Z contributed to ideas, interpreted the results and drew the some figures.

Appendix A

The hardware and software configuration of experiments is shown in Table 11 below.

See 151617

References

- [1] Zhang J, Xing W, Xing M, Sun G. Terahertz Image Detection with the Improved Faster Region-Based Convolutional Neural Network. *Sensors* (Basel Switzerland) 2018;18(7). doi: <https://doi.org/10.3390/s18072327>.
- [2] López-Tapia S, Molina R, Pérez de la Blanca N. Using machine learning to detect and localize concealed objects in passive millimeter-wave images. *Engineering Applications of Artificial Intelligence* 2018;67:81–90. doi: <https://doi.org/10.1016/j.engappai.2017.09.005>.
- [3] Thu M, Suvonvorn N. Pyramidal part-based model for partial occlusion handling in pedestrian classification. *Advances in Multimedia* 2020;2020.
- [4] Huang B, Chen R, Xu W, Zhou Q, Wang X. Improved fatigue detection using eye state recognition with hog-lbp, in: Proceedings of the 9th International Conference on Computer Engineering and Networks. Spring; 2021. p. 365–74.
- [5] Hazgui M, Ghazouani H, Barhoumi W. Genetic programming-based fusion of hog and lbp features for fully automated texture classification. *The Visual Computer* 2021;1–20.

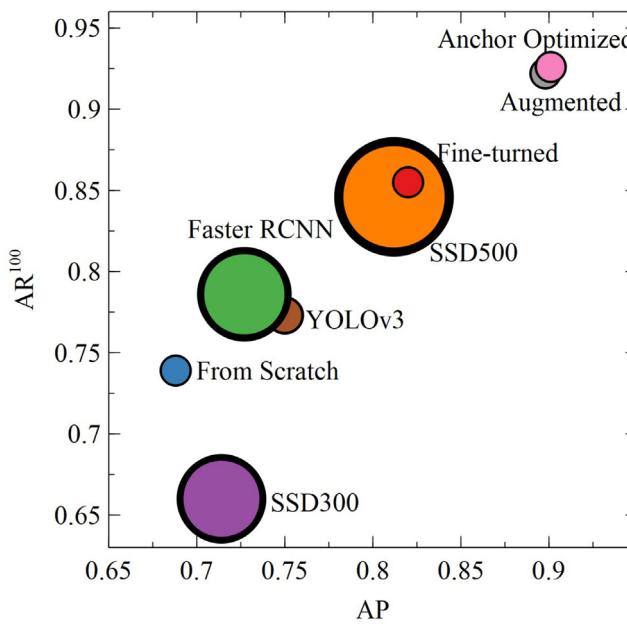


Fig. 16. Comparison with different defect detection models.

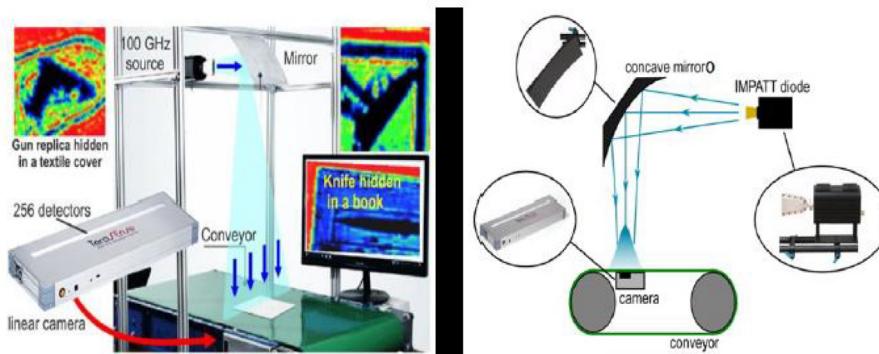


Fig. 17. Terahertz image acquisition device.

- [6] Ahmed ST, Sankar S. Investigative protocol design of layer optimized image compression in telemedicine environment. In: Procedia Computer Science 167 international Conference on Computational Intelligence and Data Science. p. 2617–22. doi: <https://doi.org/10.1016/j.procs.2020.03.323>.
- [7] Pu Y, Apel DB, Szmigiel A, Chen J. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning. *Energies* 2019;12(9):1735.
- [8] X. Wu, W. Zheng, X. Xia, D. Lo, Data quality matters: A case study on data label correctness for security bug report prediction, *IEEE Transactions on Software Engineering PP* (2021) 1–1. doi:10.1109/TSE.2021.3063727.
- [9] W. Zheng, Y. Xun, X. Wu, Z. Deng, X. Chen, Y. Sui, A comparative study of class rebalancing methods for security bug report classification, *IEEE Transactions on Reliability PP* (2021) 1–13. doi:10.1109/TR.2021.3118026.
- [10] G. Liu, Data collection in mi-assisted wireless powered underground sensor networks: Directions, recent advances, and challenges, *IEEE Communications Magazine* 59 (04 2021). doi:10.1109/MCOM.001.2000921.
- [11] Zhou Z, Lu Q, Wang Z, Huang H. Detection of Micro-Defects on Irregular Reflective Surfaces Based on Improved Faster R-CNN. *Sensors* 2019;19(22). doi: <https://doi.org/10.3390/s19225000>. <https://www.mdpi.com/1424-8220/19/22/5000>.
- [12] H. Wang, Q. Gao, H. Li, H. Wang, Y. Liping, G. Liu, A structural evolution-based anomaly detection method for generalized evolving social networks, *The Computer Journal* 65 (12 2020). doi:10.1093/comjnl/bxaa168.
- [13] Y. Shi, X. Xu, J. Xi, X. Hu, D. Hu, K. Xu, Learning to detect 3d symmetry from single-view rgbd images with weak supervision, *IEEE transactions on pattern analysis and machine intelligence PP* (06 2022). doi:10.1109/TPAMI.2022.3186876.
- [14] W. Zhou, Y. Lv, J. Lei, L. Yu, Global and local-contrast guides content-aware fusion for rgbd saliency prediction, *IEEE Transactions on Systems, Man, and Cybernetics: Systems PP* (2019) 1–9. doi:10.1109/TSMC.2019.2957386.
- [15] Penkov NV, Goltyaev MV, Astashev ME, Serov DA, Moskovskiy MN, Khort DO, Gudkov SV. The application of terahertz time-domain spectroscopy to identification of potato late blight and fusariosis. *Pathogens* 2021;10(10):1336.
- [16] Hu J, Xu Z, Li M, He Y, Sun X, Liu Y. Detection of foreign-body in milk powder processing based on terahertz imaging and spectrum. *J Infrared, Millimeter, Terahertz Waves* 2021;42(8):878–92.
- [17] Pan S, Qin B, Bi L, Zheng J, Yang R, Yang X, Li Y, Li Z. An unsupervised learning method for the detection of genetically modified crops based on terahertz spectral data analysis. *Security Commun Networks* 2021;2021.
- [18] H. Ge, M. Lv, X. Lu, Y. Jiang, G. Wu, G. Li, L. Li, Z. Li, Y. Zhang, Applications of thz spectral imaging in the detection of agricultural products, in: *Photonics*, Vol. 8, Multidisciplinary Digital Publishing Institute, 2021, p. 518.
- [19] Wang L. Terahertz imaging for breast cancer detection. *Sensors* 2021;21(19):6465.
- [20] Yin X-X, Hadjiloucas S, Zhang Y, Tian Z. Mri radiogenomics for intelligent diagnosis of breast tumors and accurate prediction of neoadjuvant chemotherapy responses-a review. *Computer Methods Programs Biomed* 2021;106510.
- [21] P. Kansal, M. Gangadharappa, A. Kumar, Terahertz e-healthcare system and intelligent spectrum sensing based on deep learning, in: *Advances in Terahertz Technology and Its Applications*, Springer, 2021, pp. 307–335.
- [22] D. Liang, F. Xue, L. Li, Active terahertz imaging dataset for concealed object detection, *arXiv preprint arXiv:2105.03677* (2021).
- [23] A.Y. Owda, N. Salmon, M. Owda, Indoor passive sensing for detecting hidden objects under clothing, in: *Emerging Imaging and Sensing Technologies for Security and Defence VI*, Vol. 11868, SPIE, 2021, pp. 87–93.
- [24] Dixit N, Mishra A. Standoff detection of metallic objects using thz waves. In: *ICOL-2019*. Springer; 2021. p. 911–4.
- [25] Xu F, Huang X, Wu Q, Zhang X, Shang Z, Zhang Y. Yolo-msfg: Toward real-time detection of concealed objects in passive terahertz images. *IEEE Sensors J* 2021.
- [26] Xie X, Lin R, Wang J, Qiu H, Xu H. Target detection of terahertz images based on improved fuzzy c-means algorithm. In: *Proceedings of 2021 Chinese Intelligent Systems Conference*. Springer; 2022. p. 761–72.
- [27] Wang T, Wang K, Zou K, Shen S, Yang Y, Zhang M, Yang Z, Liu J. Virtual unrolling technology based on terahertz computed tomography. *Optics Lasers Eng* 2022;151:106924.
- [28] Mao Q, Liu J, long Zhu Y, Lv C, Lu Y, Wei D, Yan S, Ding S, Ling D. Developing industry-level terahertz imaging resolution using mathematical model, *IEEE Transactions on Terahertz Science and Technology* 2021.
- [29] Bulgarevich DS, Talara M, Tani M, Watanabe M. Machine learning for pattern and waveform recognitions in terahertz image data. *Scientific Reports* 2021;11(1):1–8.
- [30] C. Huang, F. Jiang, Q. Huang, X.-Z. Wang, Z. Han, W.-Y. Huang, Dual-graph attention convolution network for 3-d point cloud classification, *IEEE Transactions on Neural Networks and Learning Systems PP* (2022) 1–13. doi:10.1109/TNNLS.2022.3162301.
- [31] Xie L, Zhu Y, Yin M, Wang Z, Ou D, Zheng H, Liu H, Yin G. Self-feature-based point cloud registration method with a novel convolutional siamese point net for optical measurement of blade profile. *Mech Syst Signal Process* 2022;178:. doi: <https://doi.org/10.1016/j.ymssp.2022.109243> 109243.
- [32] M. Yin, Y. Zhu, G. Yin, G. Fu, L. Xie, Deep feature interaction network for point cloud registration, with applications to optical measurement of blade profiles, *IEEE Transactions on Industrial Informatics PP* (2022) 1–10. doi:10.1109/TII.2022.3220889.
- [33] Li B, Zhang D, Shen Y. Study on terahertz spectrum analysis and recognition modeling of common agricultural diseases. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 2020;243:118820.
- [34] Antsiperov VE. Automatic target recognition for low-count terahertz images. *Computer Optics* 2016;40(5):746–51.
- [35] Zhuang Y, Chen S, Jiang N, Hu H. An effective wsenet-based similarity retrieval method of large lung ct image databases. *KSII Trans Internet Inf Syst* 2022;16(7):2359–76. doi: <https://doi.org/10.3837/tiis.2022.07.013>.
- [36] Zhuang Y, Jiang N, Xu Y. Progressive distributed and parallel similarity retrieval of large ct image sequences in mobile telemedicine networks. *Wireless Commun Mobile Comput* 2022;2022:1–13. doi: <https://doi.org/10.1155/2022/6458350>.
- [37] J. Sleiman, Terahertz imaging and spectroscopy: application to defense and security in front of the examination panel, Ph.D. thesis, Universite de Bordeaux (2016).
- [38] Morales A, Guerra R, Horstrand P, Diaz M, Jimenez A, Melian J, Lopez S, Lopez JF. A Multispectral Camera Development: From the Prototype Assembly until Its Use in a UAV System. *Sensors (Basel, Switzerland)* 2020;20. doi: <https://doi.org/10.3390/s20216129>.
- [39] Li Z, Yang Z, Song C, Li C, Peng Z, Xu W. E-eye: Hidden electronics recognition through mmwave nonlinear effects. In: *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. p. 68–81.
- [40] Liu K, Yang Z, Wei W, Bo G, Xin D, Sun C, Guoqiang G, Wu G. Novel detection approach for thermal defects: Study on its feasibility and application to vehicle cables. *High Voltage* 8 2022. doi: <https://doi.org/10.1049/hve2.12258>. n/a–n/a.
- [41] Liu S, Yang B, Wang Y, Tian J, Yin L, Zheng W. 2d/3d multimode medical image registration based on normalized cross-correlation. *Applied Sciences* 2022;12:2828. doi: <https://doi.org/10.3390/app12062828>.
- [42] Zhou W, Wang H, Wan Z. Ore image classification based on improved cnn. *Computers & Electrical Engineering* 2022;99:. doi: <https://doi.org/10.1016/j.compeleceng.2022.107819>.
- [43] Huang W, Li Y, Huang Y. Deep hybrid neural network and improved differential neuroevolution for chaotic time series prediction. *IEEE Access* 2020;8:159552–65.
- [44] Shen Y, Zhu S, Liu X, Wen S. Multistability and associative memory of neural networks with morita-like activation functions. *Neural Networks* 2021;142:162–70.
- [45] Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision* p. 2980–8.
- [46] T.-Y. Lin, P. Dollár, R.B. Girshick, K. He, B. Hariharan, S.J. Belongie, Feature Pyramid Networks for Object Detection, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, HI, USA, July 21–26, 2017, 2017, pp. 936–944. doi:10.1109/CVPR.2017.106.
- [47] Y. Liang, C. Wang, F. Li, Y. Peng, Q. Lv, Y. Yuan, Z. Huang, TFPN: Twin Feature Pyramid Networks for Object Detection, in: *31st IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2019*, Portland, OR, USA, November 4–6, 2019, 2019, pp. 1702–1707. doi:10.1109/ICTAI.2019.00251.
- [48] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mac Intell* 2017;39(6):1137–49. doi: <https://doi.org/10.1109/TPAMI.2016.2577031>.
- [49] K. He, G. Gkioxari, P. Dollár, R.B. Girshick, Mask R-CNN: Proceedings of the IEEE International Conference on Computer Vision (ICCV), CoRR abs/1703.06870 (2017).
- [50] Liu SSD. Single Shot MultiBox Detector 2016:9905.
- [51] Saleem MH, Khanchi S, Potgieter J, Arif KM. Image-Based Plant Disease Identification by Deep Learning Meta-Architectures. *Plants* (Basel, Switzerland) 2020;9(11). doi: <https://doi.org/10.3390/plants9111451>.
- [52] Widayastuti R, Yang C-K. Cat's Nose Recognition Using You Only Look Once (Yolo) and Scale-Invariant Feature Transform (SIFT). In: *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*. p. 55–6. doi: <https://doi.org/10.1109/GCCE.2018.8574870>.
- [53] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger 2017:6525.
- [54] Goodfellow I. *Adv Neural Inform Process Syst* 2014:2672.
- [55] Cao Y-J, Jia L-L, Chen Y-X, Lin N, Yang C, Zhang B, Liu Z, Li X-X, Dai H-H. Recent Advances of Generative Adversarial Networks in Computer Vision. *IEEE Access* 2019;7:14985–5006. doi: <https://doi.org/10.1109/ACCESS.2018.2886814>.
- [56] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, J. Matas, DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks, in: *CVPR 2018*, IEEE Computer Society, Los Alamitos, California, 2018. doi:10.1109/cvpr.2018.00854.
- [57] Huang X, Belongie S. Arbitrary style transfer in real-time with adaptive instance normalization. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- [58] Wang G, Giannakis GB, Chen J. Learning ReLU Networks on Linearly Separable Data: Algorithm, Optimality, and Generalization. *IEEE Trans Signal Processing* 2019;67(9):2357–70. doi: <https://doi.org/10.1109/TSP.2019.2904921>.
- [59] Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [60] Zhang X, Zou Y, Shi W. Dilated convolution neural network with LeakyReLU for environmental sound classification. In: *2017 22nd International Conference on Digital Signal Processing (DSP)*. p. 1–5. doi: <https://doi.org/10.1109/ICDSP.2017.8096153>.

- [61] Baumgartner CF, Koch LM, Tezcan KC, Ang JX, Konukoglu E. Visual feature attribution using wasserstein gans. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [62] Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [63] Shi X, Li Y. A hierarchical gamma mixture model toward hidden markov random field for high-resolution sar image segmentation. *IEEE Transactions on Geoscience and Remote Sensing* 2021.
- [64] Sridhar B. Investigations of medical image segmentation methods with inclusion mathematical morphological operations. *Treatment du Signal* 2021;38(5).
- [65] Diriker FK, Frias A, Keum KH, Lee RS. Improved accuracy of a single-slit digital sun sensor design for cubesat application using sub-pixel interpolation. *Sensors* 2021;21(4):1472.
- [66] Liu T-H, Li G-Q, Nie X-N, Wang H-J, Zhang D, Wu J-M, Liu W. Enhancement of contour smoothness by substitution of interpolated sub-pixel points for edge pixels. *IEEE Access* 2021;9:44236–46.
- [67] Chen C, Mao F, Yu J. Subpixel sampling moiré method for in-plane displacement measurement considering the symmetric errors induced by interpolation. *Applied Optics* 2021;60(5):1232–40.
- [68] Lin, Microsoft COCO: Common Objects in Context Volume 8693 (2014).
- [69] Z. Ge, S. Liu, F. Wang, Z. Li, J. Sun, Yolox: Exceeding yolo series in 2021, arXiv preprint arXiv:2107.08430 (2021).