Research on Security Inspection Method Based on Terahertz Imaging and Convolution Neural Network

JIN Bo-Yang, YAN Fang*, LIU Tong-Hua

College of Information Engineering, Inner Mongolia University of Science and Technology, Baotou, Inner Mongolia
E-mail: 0472yanfang@163.com

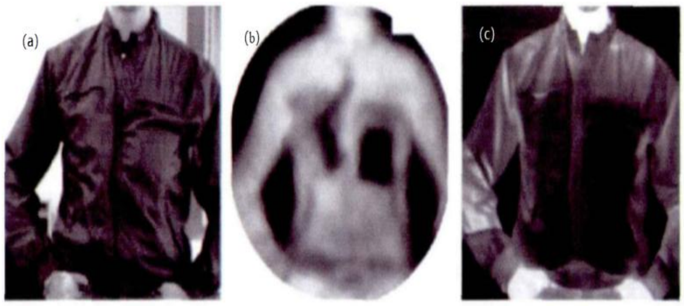
Abstract. According to the requirements of safety inspection in public places such as railway stations and airports, a dangerous goods identification method based on convolution neural network algorithm of terahertz transmission imaging technology and target detection is proposed.The research results show that the average value (MAP) of each category of AP can reach 65.02% and the average detection speed is 22.5 ms when the algorithm is trained by the self-constructed terahertz dangerous goods detection data set. In most public places the detection speed and accuracy of this method can meet the requirements of security check work , providing technical support for the research and development of security check equipment based on terahertz imaging technology.

Keywords: Terahertz imaging; Security check; Convolution neural network; Target detection

1. INTRODUCTION

In recent years, with the frequent occurrence of major crimes such as terrorist bombings and extortion cases, the task of counter-terrorism and explosion prevention has become increasingly demanding[1, 2, 3] . One of the challenges in the field of public security is how to effectively achieve rapid detection and warning of hidden weapons such as knives and guns carried by terrorists in crowded public places[4, 5, 6] . Because of the relatively high electron energy of X-rays, it is easy to cause ionization damage to the measured material, which makes the most mainstream x-ray security scanners on the market in China generally have the problem of insufficient inspection capacity. For example, the "metal gate" type of airport security screening method cannot exclude ceramic knives, explosives, flammable liquids and other non-metallic hazardous materials. Terahertz imaging-based hazardous materials detection is a target detection task that identifies the human background and hazardous materials, classifies the input image, and then determines whether the target is carrying hazardous materials[7]. The currently known dangerous goods detection algorithms are essentially based on convolutional neural networks, which are designed by designing the structure of convolutional neural networks and trained on the labeled training set to achieve the purpose of detecting the carrying of dangerous goods[8] . To address the above issues, this paper proposes a method for detecting dangerous goods targets based on terahertz imaging, constructs a dangerous goods carry detection dataset, adapts to the dangerous goods carry detection task by normalizing the terahertz images and adjusting the parameters of the original YOLO-V4 network, and investigates the effect of improving the model to further improve the model recognition accuracy and enhance the robustness of the model.

Research in the THz band has expanded the ability of humans to "see". A comparison of the effects of visible light, terahertz and infrared imaging is shown in Figure 1. In comparison to visible light, terahertz waves can effectively penetrate a wide variety of materials such as knitted clothing, paper advertising images, and metal plastics. Infrared and optical imaging have disadvantages such as low resolution and high influence by the environment. Compared with X-rays, terahertz waves are lower in energy and do not produce ionizing radiation. Based on the above characteristics, terahertz imaging technology in the field of human security screening has received more and more attention. .



(a) Visible light imaging; (b) Terahertz imaging; (c) Infrared imaging

Fig. 1 Comparison of imaging results and effects[9]

Table 1 Comparison of Various Wave Data

Types of waves	Frequency THz	Energy eV	Can it penetrate cloth and paper
Visible light	380~750THz	1.64~3.19eV	Can
Infrared ray	10THz~380THz	0.416~1.64eV	Incomplete
Terahertz	0.1~10THz	0.413~416meV	Can
X-ray	>3*10^4 THz	124eV~1.24MeV	Can

2. DEVICE AND IMAGING EXPERIMENTS

2.1. Experimental Device

The experimental instrument is a BRUKER TPS spectra 1000 terahertz pulse time domain spectrometer manufactured by Bruker, Germany. The transmission THz-TDS system mainly consists of a femtosecond laser, a terahertz detection system, a terahertz radiation generation device and a time delay control system, and its optical path schematic is shown in Figure 2. In the terahertz pulse spectroscopy system, the high frequency pulses emitted from the femtosecond laser are divided into pump pulses and detection pulses after passing through the beam splitter. The pump pulse is incident on the terahertz radiation generation device after the time delay system to generate terahertz pulses, and the emitted terahertz pulses pass through the sample and are co-linearly incident on the terahertz detection device with the detection pulses. The time delay between the pump pulse and the detection pulse is adjusted by controlling the time delay system, and the entire time domain waveform of the terahertz pulse can be detected. The time domain spectrum of the terahertz pulse can be converted into frequency domain spectrum by Fourier transform to obtain optical parameters such as refractive index and absorption coefficient. After computer processing, the absorption of the substance to terahertz is plotted as a spectrogram, and then the spectrogram is analyzed to obtain useful information in it and determine the structural composition of the sample under test.

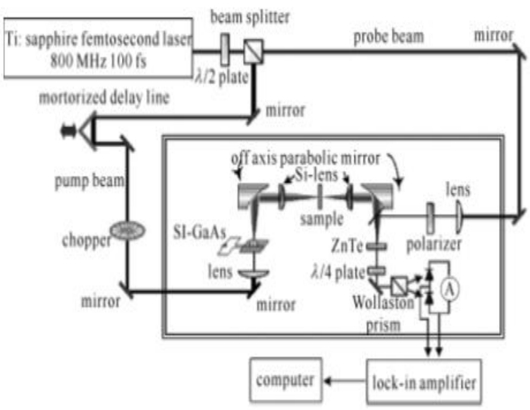


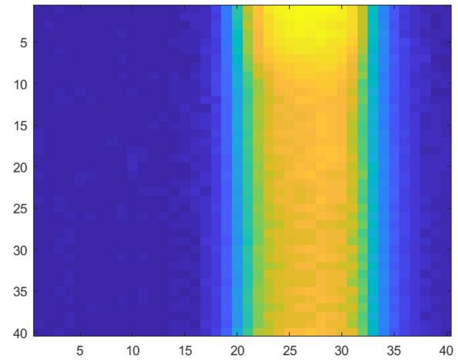
Fig. 2 Terahertz Continuous Wave Point Scanning Imaging System[10]

2.2. Terahertz Hazardous Materials Imaging Experiment

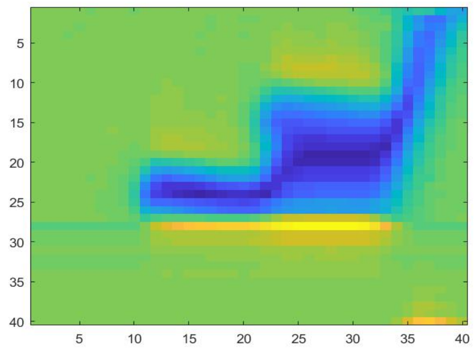
The scanning step of this experimental panning table is 500μm.The maximum scanning range of the sample stage was 1.95 cm × 1.95 cm, and the number of scanned pixels was 1600, and the imaging time was 58 minutes. A vacuum pump was started to evacuate the sample cavity to remove the influence of various components in the air on the terahertz signal. After scanning, the time domain spectrum, frequency domain spectrum, transmission spectrum and absorption coefficient are saved for subsequent data processing.

First, the ceramic blade was scanned and imaged using the terahertz imaging system , considering that the suspect carrying dangerous goods will not leak the dangerous goods directly outside the body for security check, so the experiment used paper and fabric to cover for scanning and imaging. Figure 3(a) shows the imaging effect of Ceramic blade sample at 5.97 THz, and the outline of the ceramic blade can be clearly seen in the scanning image. The scanning imaging result consists of yellow, blue and green parts, where the yellow area is the low transmission of terahertz waves by the ceramic blade sample, and the blue area is the part where terahertz penetrates the scanned sample. The transition area between the ceramic knife and the air is the green area.

Secondly, this experiment selected a thin aluminum sheet cut into the shape of a gun as a scan sample, the sample size is 1.3cm×1.0cm, Figure 3(b) shows the effect of the thin aluminum gun after imaging, from the scan imaging image can clearly see the thin aluminum gun outline. The scanned image results consist of green and blue color, the blue area is the low transmission rate of terahertz wave for the thin aluminum gun sample, and the green area is the part of terahertz penetration of the scanned sample.



(a) Imaging effect of Ceramic blade sample



(b) Imaging effect of thin aluminum gun sample

Fig.3 Experimental results of terahertz imaging of hazardous materials

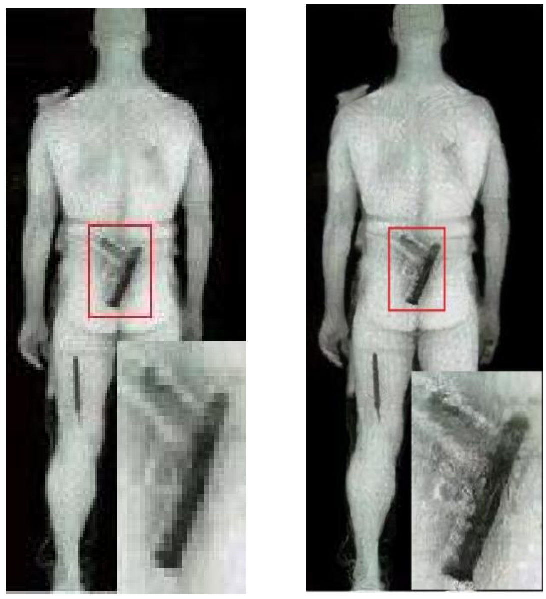
3. HAZARDOUS GOODS TARGET DETECTION BASED ON YOLOV4 ALGORITHM

3.1. Effect enhancement using SRGAN dataset

From the above experimental results, it can be seen that the terahertz security images have the disadvantages of single tone, low resolution, poor clarity and contrast. If the terahertz security images are directly fed into the target detection network for training and testing, it is not possible to achieve better detection results. Therefore, image pre-processing is needed to improve the above-mentioned problems in terahertz security images. Directly processing the terahertz security images by traditional preprocessing methods, such as grayscale conversion, image interpolation, image filtering, image denoising, etc, cannot improve the quality of terahertz security images significantly. By performing super-resolution operation on the low-quality images can be converted to high-resolution images, which can well solve the problem of unclear low-quality images.

In this paper, we combine super-resolution generative adversarial network SRGAN to perform pre-processing operations on terahertz security images to transform low-resolution terahertz security images into high-resolution images, obtaining higher visual quality as well as more realistic and natural texture features.The network structure of SRGAN is divided into two parts: generator and discriminator, and Figure 4 shows the SRGAN discriminator model. The generator uses all the same dense residual blocks of SRResNet as the basic unit, and the structure contains a "residual-re-residual" structure, which allows the residuals to be used at different levels and ensures that the gradient information can be effectively transferred, thus enhancing the robustness of the generative adversarial network. Also the use of residual scaling and smaller initialization helps to train deeper networks, thus improving the perceptual quality and further improving the repaired texture.

layers are too deep[12]. the CSPnet construction is more refined, where the backbone part still performs the stacking of residual blocks; the other part is sent to the end of the network like a residual edge, after a brief processing, and its is the stacking of residual blocks of the original YOLO-V3 divided into two.



(a) Low resolution image[11] (b) SRGAN reconstructed image

Fig. 5 Experimental results after super-resolution reconstruction.

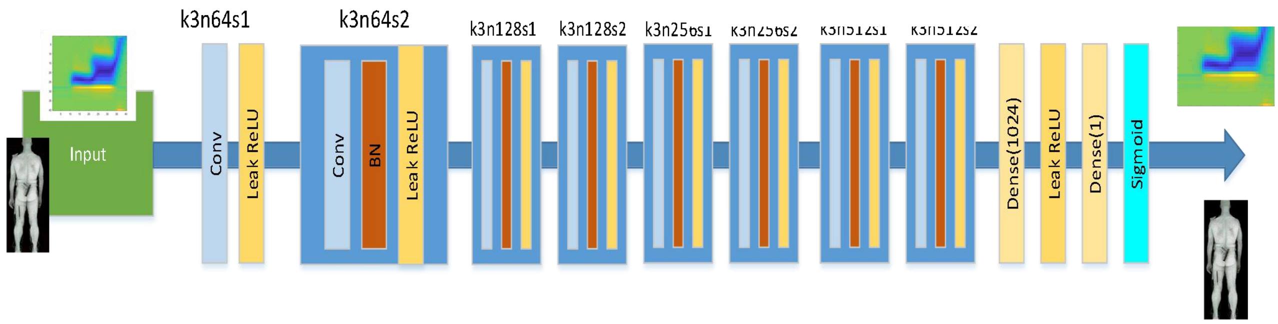


Fig. 4 SRGAN discriminator model

Figure 5 shows the results of the image super-resolution reconstruction with a 4× magnification factor. The low-resolution image is shown in Fig. (a), and the super-resolution image after the experiment with the method used in this design is represented in Fig. (b). The specific position of the image after partial magnification is marked with a red box, and the effect of the magnified image is shown in the lower left corner of the image. Comparing Fig. (a) as well as Fig. (b), it can be seen that the image obtained after super-resolution processing of the original terahertz security image by SRGAN becomes clearer in detail texture, more complete information, sharper and clearer edge contours, and better visual effect of the image.

3.2. CSPDarknet-53 feature network extraction

CSPDarknet-53 emulates the thinking of deep residual networks by complementing a series of residual network compositions to avoid gradient explosion when the network

3.3. Activation function

The network structure in YOLO-V3 is Darketnet53, which consists of a series of residual networks. In Darket53, there is a Resbolck-body module, which consists of a single downsampling and multiple residual structure stacking, while inYOLO-V4 , it makes modifications to this part, it is to LeakyReLU modify DarketConv2D the activation function from Mish, and the convolution block from DarketConv2D_BN_Leaky into DarketConv2D_BN_Mish The activation function equation is as follows.

$$Mish = x \times \tanh(\ln(1+\exp(x))) \quad (1)$$

X is the output of the upper level node and $Mish$ is the input of the lower lower level node .

Mish's equation and image of the function are as figure 6.

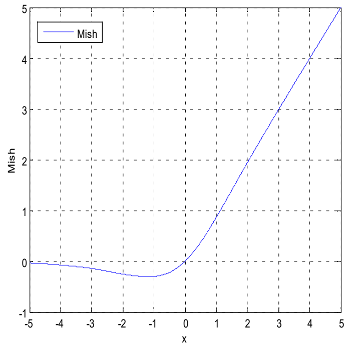


Fig. 6 Mish Activation Function

3.4. Experimental analysis of terahertz hazardous material target detection

The terahertz hazmat image dataset in this paper is obtained from 98 images acquired and collected by ourselves with various degrees of data expansion. The experimentally acquired images are rotated, symmetrical, flipped and the sensitivity of the images is changed to improve the robustness of the model. After selection, the dataset consists of 1274 terahertz hazmat images. The target annotation is carried out by Label lmg image annotation tool, and the annotated data are divided into two categories, which are firearms and knives. The standard information of each image after annotation includes the geometric center of the annotation frame and the length and width of the annotation frame . The training data include 1019 sets of training sets and 255 sets of validation sets.

The experiments were conducted using Window 10 operating system, 4GB video memory 1650Ti GPU, 32GB CPU to complete the computing session, and tensorflow deep learning framework. The training model parameters were set as shown in Table 2. After training, the loss value of each iteration is read from the log file.

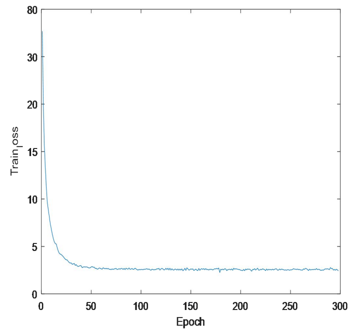
Table 2 Training Model Parameter Setting

Parameter Name	Parameter value
Batch	10
Epochs	300
Basic learning rate	0.001
Decay rate of learning rate	10%
Step Size	66

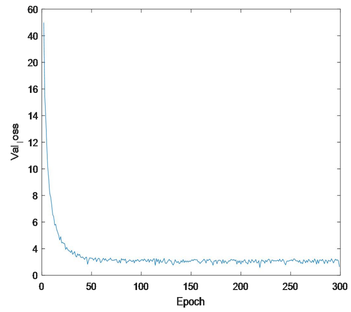
Figure 7 shows the loss function curves of the training set and the test set generated by the training process, the vertical coordinate is the loss value and the horizontal coordinate is the number of training steps. The loss function value reaches 744 and 56 at the beginning of the training process, and the loss value slowly starts to decrease as the number of training iterations increases. After 50 iterations of training, the loss value was 2.6 and 3.3, and as the batches increased, the loss still decreased, but the rate of decrease began to slow down.The training was carried out up to about 250 times, and the Loss curve values were basically stable with no significant trend and converged to 2.5 and 3.0.

Since only the precision and recall values are used to judge the performance of the model with a certain degree of chance, this paper adds the index of MAP value to comprehensively evaluate the effectiveness of the model.

The AP refers to precision and recall when using different confidence levels. The MAP refers to the integrated area of the curve plotted using the combination of points with different confidence levels. The MAP of this model is defined as the average of the AP value of the gun and the AP value of the tool. The MAP histogram before and after model reconstruction is shown in Figure 8.



(a) Training set loss



(b) Test set loss

Fig. 7 Loss Curve of Training Model.

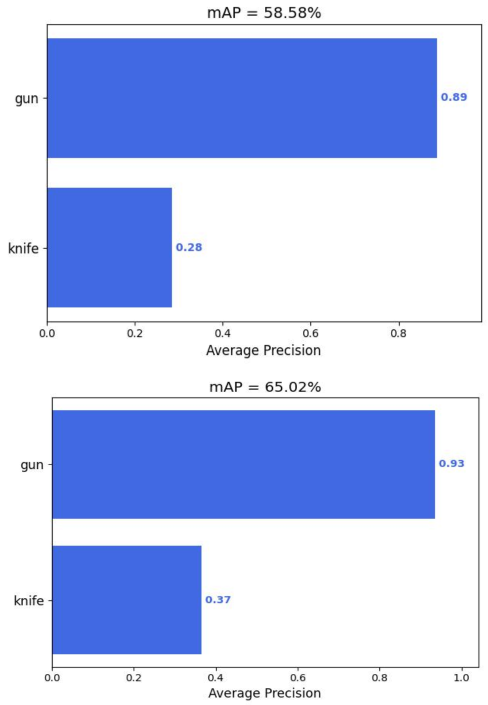


Fig. 8 AP Curve of Training Model

In this paper, we use the TP value to indicate the number of targets that correctly detect the dangerous goods,FP value to indicate the he number of targets that cannot correctly detect the dangerous goods, and FN value to indicate the number of targets that miss the detection of the dangerous goods.

The Precision and Recall are shown below (2).(3)

$$precision = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

As seen in Table 3, the MAP value of SRGAN-YOLO-V4 is 6.44 percentage points higher than that of YOLO-V4, and the detection speed of each target is improved by 3.2ms.

Table 3 Performance Comparison of SRGAN-YOLO-V4 and YOLO-V4

Detection algorithm	MAP/%	Detection Speed ms
SRGAN-YOLO-V4	65.02%	22.5ms
YOLO-V4	58.58%	25.7ms

The detection effect of the test set of pictures is shown in Figure 9. For carrying dangerous goods, pink and green boxes are used to mark them respectively, and the words gun and knife class are added, along with confidence levels. The experiment proves that the detection speed as well as the detection accuracy of the method can meet the needs of most public security screening work.

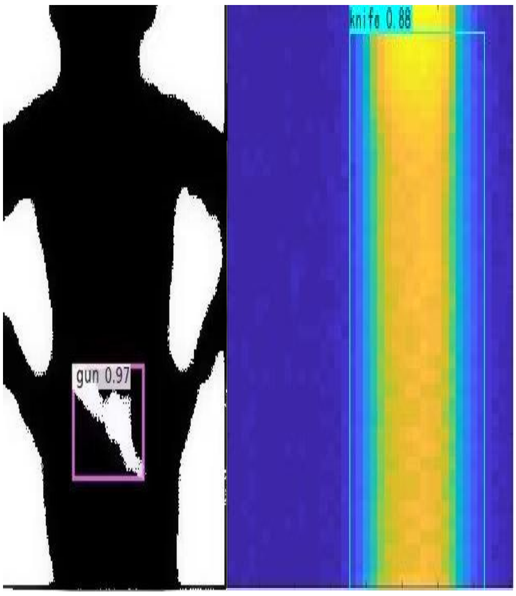


Fig.9 Schematic Diagram of Terahertz Dangerous Goods Security Inspection Results

4. Conclusion

In this paper, a super-resolution reconstruction of terahertz images of hazardous materials based on SRGAN algorithm is performed to enhance the dataset. By comparing the results of the YOLO-V4 algorithm with those of the SRGAN super-resolution reconstruction, it can be seen that the YOLO-V4 algorithm with SRGAN reconstruction has obvious advantages in terms of MAP value and detection speed. Based on this, the yolov4 deep convolutional network was used to train the terahertz detection image dataset with dangerous goods, and good security results were obtained. The detection quality can be further improved by improving the activation function and enhancing the small target localization capability.

5.ACKNOWLEDGEMENTS:

We thank the National University of Defense Technology for providing the experimental platform for terahertz spectroscopy, and Professor Dongwen Zhang of the College of Science, National University of Defense Technology for his help and helpful discussions on the experiments.

6.REFERENCES:

[1] Jiang L H, Wang W S, Tong Huixin, et al. Research P rogress of Terahertz Imaging Technology in Human Security [J]. Journal of Shanghai University of Techn ology, 2019, 41 (01): 46-51.

[2] Yan Z X, Lin X, Tong H X, et al. Automatic identific ation and segmentation of suspicious objects in terah ertz images [J]. Journal of Shanghai University of Tec hnology, 2018, 40 (05): 435-443.

[3] Zhou J S, Zhou D L, Yuan Y, et al. Design of a quasi-optical system for terahertz imaging [J]. Electronic M easurement Technology, 2018, 41 (10): 142-146.

[4] Gu Q, Sang W, Jiang J Q. Station open millimeter wa ve/terahertz human security checktechnology and d evelopment trend [J]. Police Technology, 2016 (03): 7-10.

[5] Cheng B B, Li H P, et al. Application of terahertz ima ging technology in station open security check [J]. Jo urnal of Terahertz Science and Electronic Information, 2015, 13 (06): 843-848.

[6] Li Q, Ding S H, et al. Continuous terahertz reflection scanning imaging experiment of hidden objects [J]. C hina Laser, 2012, 39 (08): 206-211.

[7] Wang B, Wang X K, Yu Y, et al. Terahertz linear arra y fast scanning imaging [J]. China Laser, 2019, 46 (06): 280-285.

[8] Redmon J,Divvala S,Girshick R,et al. You only look once:Unified,real-time object detection[C]/ /Pr Procee dings of the IEEE conference on computer vision and pattern recognition.2016:779-788.

[9] May T. Safe VISITOR: visible, infrared, and terahertz object recognition for security screening Internatio nal Society for Optics and Photonics, 2009:73090E-79090E-8.

[10] Wu, Ronglin, Shao, Zhengzheng, Chang, Shengli, Zh ang, Xueqi, Li, Heping, Li, Xinhua. Raman spectrosc opic study of multi-walled carbon nanotubes with diff erent parameters[J]. Spectroscopy and Spectral Analy sis,2014,34(04):982-985.

[11] Li Zeng. Research on terahertz imaging system and su per-resolution image reconstruction algorithm [D]. Zh ejiang University,2018.

[12] HE K, ZHANG X, REN S, et al. Deep residual learni ng forimage recognition[C]//IEEE Conference on Co mputer Visionand Pattern Recognition, 2016: 770-77 8.