2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

A Defect Recognition and Classification Method Based on Improved Convolutional Neural Network and Terahertz Time-Domain Spectroscopy System

Liu Yang[1], Xiuwei Yang[2], Teng Li[1], Aoyu Zhu[1] and Jinhong Li[1(✉)]

[1] School of Mathematics and Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
[2] Faculty of Electronic Information and Control Engineering, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
lijinhong@qlu.edu.cn

Abstract. Convolutional neural networks (CNN) can perform defect recognition and classification, saving time compared to traditional methods. However, traditional CNN are difficult to achieve accurate differentiation due to insufficient feature extraction capability and low computational efficiency when dealing with scenes with complex backgrounds and similar defect categories. To solve these problems, this paper proposes an improved CNN based on multimodal data fusion to achieve efficient automated defect recognition and classification by combining the technical advantages of terahertz time-domain spectral system. Firstly, the spectral data of the samples are obtained by a terahertz time-domain spectroscopy system, and the pre-processed spectral data are imaged. Second, the absorption coefficients were obtained by building a terahertz propagation model inversion. Then, the terahertz absorption coefficients are deeply fused with the image data to construct a multimodal dataset as the network input. Convolutional blocks with multi-layer asymmetric convolutional kernels are designed in the convolutional layer to enhance the accuracy and classification speed of defect recognition by strengthening the feature extraction and learning capabilities. Meanwhile, jump connections are chosen between the convolutional blocks, aiming to resist the problems of gradient vanishing and overfitting. Numerical experiments demonstrate that the improved CNN attains an accuracy of 99.4% in defect classification, with an F1 score of 0.99, and 100% accuracy in the confusion matrix validation set. Compared with traditional CNN, the accuracy is increased by 6% and the F1 score is improved by 4%. This provides reliable technical support for defect recognition and classification in complex scenes.

Keywords: Convolutional Neural Network, Defect detection and classification, Terahertz Time-Domain Spectroscopy System.

1 Introduction

In the contemporary textile industry, fabrics serve as a crucial raw material, playing a pivotal role in various sectors, including garment manufacturing and home decoration.

Nevertheless, during the production and transportation of fabrics, defects such as snags, holes, stains, twists, yarn piles, and abrasions frequently occur within or on the fabric surface. These issues arise due to factors like equipment malfunctions and improper process control. Such defects not only compromise the aesthetic appeal of the finished products but also significantly affect the durability and stability of the fabrics. Additionally, they increase the difficulty of cutting operations, thereby reducing production efficiency and product quality. Consequently, the development of efficient and accurate methods for detecting and classifying fabric defects has become a critical concern in the textile industry.

Traditional fabric defect detection predominantly relies on manual visual inspection. This approach is not only time-consuming and labor-intensive but also highly susceptible to subjective influences, which can undermine the accuracy of the detection process [1]. With the rapid advancement of computer vision and artificial intelligence technologies, numerous non-destructive testing methods for fabric defects have emerged. These include X-ray detection, ultrasonic detection, infrared detection, and machine-vision-based detection methods [2]. Among them, although X-ray technology can penetrate various types of fabrics, the associated equipment is costly, and it involves radiation risks [3]. Ultrasonic detection can assess the internal structural integrity of fabrics, yet it is significantly influenced by material properties, making it challenging to detect defects in complex-shaped fabrics [4]. Infrared detection, leveraging thermal imaging technology, can quickly locate abnormal regions. However, it is highly sensitive to the surface roughness of materials and ambient temperature variations [5]. Machine-vision-based detection methods can clearly capture the fine textures and surface defects of fabrics. Still, they impose stringent requirements on light stability, and changes in ambient light can readily affect the detection accuracy [6]. Evidently, these existing techniques have certain limitations when applied to fabric defect detection.

In recent years, the terahertz time-domain spectroscopy (THz-TDS) system has shown great promise in the textile industry's non-destructive testing applications. This system benefits from its strong penetration ability, high resolution, insensitivity to light interference, and fingerprint spectroscopy characteristics [7]. It can simultaneously detect multiple types of defects both inside and on the surface of fabrics. In fabric inspection, the THz-TDS system can identify fabric structures and compositions without causing damage to the materials [8], providing a novel technical approach for fabric defect detection. However, in practical applications, to further enhance recognition accuracy and classification speed, the automatic recognition and classification of fabric images containing defects are essential. Most previous research methods focused on likelihood-based or feature-based classification [9]. Nevertheless, these methods encounter challenges in practical applications, such as limited computational capabilities or insufficient experience in signal feature engineering, which hinder their ability to meet the requirements of automatic detection in industrial inspection. To address these issues, fabric defect classification methods based on image processing and deep learning have gradually become a research hotspot [10]. Among these, the Convolutional Neural Network (CNN) has emerged as one of the mainstream techniques for solving fabric defect classification problems. CNN's ability to automatically extract features at different levels from data and its excellent depth scalability contribute to its popularity [11]. For

instance, Wang et al. developed a terahertz characterization system based on a deformable attention CNN framework to achieve automatic defect localization [12]. However, it exhibits poor accuracy in identifying certain subtle or complex defect features. Kim et al. analyzed tiny internal defects in materials using a deep learning model of terahertz signals based on CNN [13], but the method has high computational complexity and is difficult to implement in real-time applications. Zhang et al. proposed a defect detection model based on a hybrid attention multiscale non-jumping U-type deep convolutional self-encoder [14]. Still, as the number of network layers increases, it is prone to problems such as gradient vanishing and information loss. Mengiste et al. explored the potential of combining deep transfer learning architectures to improve the diagnostic accuracy of automatic identification [15].

Faced with complex fabric backgrounds, easily confusable defect categories, and blurred fabric edges, traditional CNN algorithms with limited feature extraction capabilities and long running times struggle to efficiently distinguish between various types of defects. To address this, this paper integrates the advantages of the THz-TDS system and the CNN model and proposes an improved CNN method for defect recognition and classification. This method first employs a THz-TDS system to acquire the spectral signals of samples and then images the pre-processed spectral signals. Subsequently, a sample transmission model is established and inverted to calculate the terahertz absorption coefficient, verifying the feasibility of the THz-TDS system in fabric defect identification. The terahertz absorption coefficients and image data are then correspondingly fused and input as a dataset into the network. In the model design, multi-layer convolutional blocks with asymmetric convolutional kernels and a skip connection structure are utilized to enhance the model's recognition accuracy and generalization ability. Experimental results demonstrate that the improved CNN model can better adapt to the diversity and complexity of sample defects. It performs well in terms of the confusion matrix, accuracy, precision, recall, and F1 score, enabling accurate recognition and classification of different fabric defects.

The paper is organised as follows. Section 2 describes the terahertz propagation model in fabrics. The proposed improved CNN network is described in Section 3. Section 4 verifies that the proposed network is efficient through numerical experiments. Section 5 presents the conclusions.

2 Terahertz propagation model

It is assumed that the terahertz wave emitted by the THz-TDS system in transmission mode is a plane wave. The incident terahertz wave is called $E_{THz}(w)$, and when the terahertz wave passes directly through the air without considering atmospheric losses and scattering effects, the received signal is called $E_{ref}(w)$ reference wave. When the terahertz wave is incident from one medium to another, the received signal is called the sample wave $E_{sam}(w)$, as shown in Fig. 1.

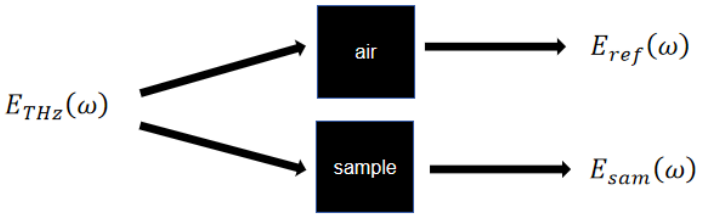


Fig. 1. Schematic diagram of transmissive terahertz wave transmission.A figure caption is always placed below the illustration.

The interaction between the terahertz wave and the sample is described by the Maxwell's system of equations [16], the connection between the electromagnetic field quantities is established according to their boundary conditions, and the reflection coefficients (r_{as}) and the transmission coefficients (t_{as}) at the different dielectric demarcation surfaces are obtained through Fresnel's formula [17,18].

$$r_{as} = \frac{\tilde{n}_s(\omega)\cos\varphi_a - \tilde{n}_a(\omega)\cos\varphi_s}{\tilde{n}_s(\omega)\cos\varphi_a + \tilde{n}_a(\omega)\cos\varphi_s}, \tag{1}$$

$$t_{as} = \frac{2\tilde{n}_a(\omega)\cos\varphi_s}{\tilde{n}_s(\omega)\cos\varphi_a + \tilde{n}_a(\omega)\cos\varphi_s}, \tag{2}$$

where $\tilde{n}_a(\omega)$ denotes the negative refractive index of terahertz waves in air, $\tilde{n}_s(\omega)$ denotes the complex refractive index of terahertz waves in the sample, φ_a and φ_s denote the angles of incidence and refraction, respectively.

As a terahertz wave propagates through a sample, its amplitude decays and its phase is delayed, and the change is described by the propagation factor [19],

$$p(\omega, L) = \exp\left(\frac{-j\tilde{n}(\omega)\omega L}{c}\right), \tag{3}$$

where $\tilde{n}(\omega)$ is the complex refractive index, ω is the angular frequency, L is the propagation distance, and c is the speed of light.

Assuming that the total path of terahertz propagation in air is Z when the sample is not placed, the projected distance in the direction of fibre propagation is h, d is the thickness of the sample, and the length of a single reflected ray is x.

When the sample is placed and the terahertz wave is incident on the sample at a certain angle, the Fabry-Perot effect (F-P effect) [20,21] occurs, where part of the wave is reflected and part of the wave penetrates through the sample and refracts and reflects

several times internally, and the terahertz wave propagates through the air over a distance of $Z-h$,obtaining the Eq. 4.

$$\begin{cases} E_{ref}(\omega)=E_0(\omega)\,p_{air}(\omega,L), \\ E_{sam}(\omega)=E_{t0}(\omega)+\sum_{k=1}^{\infty}E_{tk}(\omega), \\ \tilde{n}(\omega)=n(\omega)-ik(\omega), \\ h=d\cos\varphi_1, \\ x=\dfrac{d}{\cos\varphi_2}, \end{cases} \tag{4}$$

The reference and sample signals are obtained from Eqs. 1-4 as shown in Eq. 5 and Eq. 6.

$$E_{ref}(\omega)=E_{THz}(\omega)\exp\left(\frac{-j\tilde{n}_a(\omega)\omega L}{c}\right), \tag{5}$$

$$\begin{aligned} E_{sam}(\omega)&=E_{THz}(\omega)\,p_{air}(\omega,Z-h)t_{as}p_s(\omega,x)t_{sa} \\ &+E_{THz}(\omega)\,p_{air}(\omega,Z-h)t_{as}p_s(\omega,x)t_{sa}\sum_{k=1}^{\infty}r_{sa}^2\,{p_s}^2(\omega,x)^k, \end{aligned} \tag{6}$$

Therefore, the transfer function is obtained as in Eq. 7.

$$\begin{aligned} H(\omega)&=\frac{E_{sam}(\omega)}{E_{ref}(\omega)} \\ &=\frac{4\tilde{n}_a(\omega)\tilde{n}_s(\omega)\cos\varphi_a\cos\varphi_s}{\left[\tilde{n}_a(\omega)\cos\varphi_s+\tilde{n}_s(\omega)\cos\varphi_a\right]^2}\exp\left\{\frac{-j\left[x\tilde{n}_s(\omega)-h\tilde{n}_a(\omega)\right]\omega}{c}\right\}, \\ &\cdot\left(1+\sum_{k=1}^{\infty}r_{sa}^2p_s^2(\omega,x)^k\right) \end{aligned} \tag{7}$$

The F-P effect factor was obtained as

$$FP(\omega)=1+\sum_{k=1}^{\infty}r_{sa}^2p_s^2(\omega,x)^k, \tag{8}$$

In order to extract the correct target signal, assume that the terahertz is incident on the sample vertically and the F-P effect is 1, $\cos\varphi_a=\cos\varphi_a=1$, $\tilde{n}_a=1$, to obtain its transfer function as

$$H_v(\omega)=\frac{4\tilde{n}_s(\omega)}{\left[1+\tilde{n}_s(\omega)\right]^2}\exp\left\{\frac{-j\left[\tilde{n}_s(\omega)-1\right]\omega d}{c}\right\}, \tag{9}$$

In the case of weak absorption, when $n(\omega)\geq k(\omega)$, it can be neglected $k(\omega)$. From the relation between the mode and the supplementary angle and Lambert's law

$$\begin{cases} H(\omega)=\rho(\omega)e^{-j\phi(\omega)}, \\ \rho(\omega)=\left|H(\omega)\right|, \\ \phi(\omega)=\arg\left[H(\omega)\right], \\ n_s(\omega)=\dfrac{\phi(\omega)c}{\omega d}, \\ k_s(\omega)=\dfrac{c}{\omega d}\left\{\ln\dfrac{4n(\omega)}{\rho(\omega)\left[n(\omega)+1\right]^2}\right\}, \\ \alpha(\omega)=\dfrac{2\omega k(\omega)}{c}, \end{cases} \tag{10}$$

Obtaining the absorption coefficient

$$\alpha(\omega)=\frac{2}{d}\left\{\ln\frac{4n_s(\omega)}{\rho(\omega)\left[n_s(\omega)+1\right]^2}\right\}, \tag{11}$$

where the real refractive index of the sample is $n_s(\omega)$, the extinction coefficient is $k_s(\omega)$, and d is the thickness of the sample.

Therefore, the strength of the sample's ability to absorb terahertz waves can be learnt based on the magnitude of the absorption coefficient, thus distinguishing different defects of the sample.

3 Improved CNN algorithm

The CNN is regarded as a deep-learning architecture [22]. Due to the slow classification speed caused by the massive data measured by the terahertz time-domain spectroscopy system, traditional CNN perform poorly in identifying and classifying defects based on single defective features. Thus, improving the CNN to optimize and innovate in terms of classification accuracy and speed is necessary. Therefore, this section presents an efficient CNN architecture for classification, as illustrated in Fig. 2.

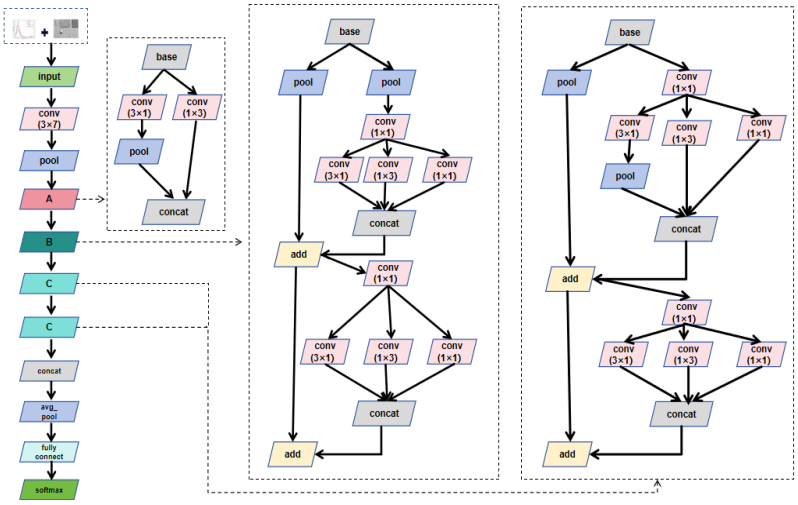


Fig. 2. Architecture of the improved CNN algorithm.

3.1 Input Layer Improvement

The input layer serves as the starting point for neural networks to process data. It receives input data and forwards it to the convolutional layer for feature extraction. When the input layer only processes a single piece of data or image, the model's information acquisition is relatively limited. This limitation can result in poor performance when handling complex tasks, as the model fails to analyze problems from multiple perspectives or dimensions. To address this issue, this paper transforms the original complex one-dimensional spectral data into a two-dimensional matrix similar to image data. Then, it corresponds one-to-one with image data in the channel dimension, fusing and feeding it into the network as input data. Specifically, each signal frame in the dataset is organized into a 2D array with a size of 2×1024.

3.2 Improvement of convolutional layers

The convolutional layer is a multi-layer convolutional block. Asymmetric convolutional kernels are used, and jump connections are introduced between these blocks. Inside the convolutional layer, the components are connected in the following order: the first convolutional block, the pooling layer, feature fusion module A, information enhancement module B, and two depth processing modules C, as depicted in Fig. 2. Input data passes through the first convolution block (convolution2dLayer) to extract general features, generating the first feature map. The first convolution block has a 3×7 convolution kernel and reduces the spatial dimension in steps of (1, 2).

To minimize the number of trainable parameters while maintaining feature extraction quality, the feature fusion module A uses two juxtaposed layers of asymmetric convolution kernels: a 3×1 vertical kernel (the second convolution block) and a 1×3

horizontal kernel (the third convolution block). Notably, both convolution and pooling operations use steps of (1, 2), which efficiently reduces the horizontal dimension of the feature map and cuts computational costs. The second feature map output by module A is fed into the maximum pooling layer. Through downsampling, it produces the third feature map, facilitating more effective learning of discriminative features.

The information enhancement module B contains two asymmetric convolutional modules. In the first module, the fourth convolutional block connects to 3×1, 1×3, and 1×1 asymmetric convolution kernels and concatenates them along the depth dimension using the Depth-concatenation Layer at the connection layer. The fourth convolutional block directly following the pooling layer output extracts features, while the one juxtaposed with the second and third convolutional layers reduces the channel dimension. The features from the first asymmetric convolutional block are fused with the maxpooled features of the third branch in the ADDITIONLAYER. This significantly enhances the overall feature representation compared to traditional CNNs' single-feature extraction paths. The fused features then enter the second asymmetric convolution module, undergoing 1×1 convolution, followed by parallel 3×1, 1×3, and 1×1 convolutions. The output feature maps are concatenated in depth by the Depth-concatenation Layer, and then element-wise fused with the third-branch output by the addition Layer to gradually learn feature information.

The depth processing module C consists of two asymmetric convolution modules. The first module first goes through a 1×1 convolution layer, then juxtaposes 3×1, 1×3, and 1×1 asymmetric convolution kernels. The 3×1 kernel connects to the maxpooling layer, reducing feature dimensionality and computation. Then, 1×3 and 1×1 convolution kernels are used for fusion with the Depth-concatenation Layer. The fused features are combined with the fifth-branch pooled features via the addition layer, integrating features from different dimensions to enrich feature information. The features fused by the addition layer are input to the second asymmetric convolution module, which first applies a 1×1 convolution layer, then parallel 3×1, 1×3, and 1×1 asymmetric convolution kernels, and concatenates the output along the depth dimension at the connection layer. Finally, the concatenated features are element-wise fused with the fifth-branch output by the addition Layer to obtain the final feature map.

3.3 Improvements in the connectivity layer

To boost the classification model's accuracy and alleviate the gradient vanishing problem caused by the activation function in the network, Skip Connections are added between convolutional blocks. These connections preserve the information of features extracted in each block throughout the network and combine them, enriching the classification model. Importantly, Skip Connections enable modules to learn residual information instead of traditional true outputs. Consequently, the improved CNN model can better resist gradient vanishing and overfitting during network training.

At the end of the model, the feature maps from the last depth processing module C are combined with the previous jump link output via the connection layer. Then, a global averagePooling2dLayer converts the feature maps into fixed-length vectors, which are fed into the Fully Connected Layer. A Dropout Layer is added to prevent

overfitting and ensure the model's generalization ability. The number of hidden units in the fully connected layer, denoted as C, is set to match the number of signal modulation method types in the dataset. Finally, a Softmax Layer is connected to perform the multi-category classification task and output defect classification results. The detailed layers of the improved CNN architecture are shown in Table 1.

Table 1. Architecture details diagram.

Layer	Output Volume	Dropout
input	2x1024x1	-
conv	2x512x64	0.1
pool	2x256x64	-
A	2x128x64	0.1
B	2x32x128	0.1
C	2x16x128	0.1
C	2x8x128	0.1
concat	2x8x384	-
Pool	1x1x384	-
classifcation	1x1xC	0.1

3.4 Model training strategy

The improved CNN model is trained with the data from the training set. The stochastic gradient descent optimization algorithm, combined with the cross-entropy loss function, is employed to dynamically adjust the model's weight parameters. After 4500 training iterations, the loss function value of the model on the training set converges to a stable low value.

When using stochastic gradient descent for model training, 128 samples are utilized for parameter updates in each iteration to strike a balance between training efficiency and memory consumption. The maximum number of training epochs is set to 60, meaning the model will undergo 60 full iterations over the entire training set. The initial learning rate is set at 0.01, and a segmented learning rate adjustment strategy is adopted. Specifically, the learning rate decays to 0.1 times its original value every 30 epochs. This approach allows for the gradual fine-tuning of model parameters as training progresses, helping to avoid getting trapped in local optimal solutions.

Before each training epoch, data shuffling is enabled. This ensures that the model encounters data in different orders during various training cycles, thereby enhancing its generalization ability. The model's performance is monitored using the validation set, with the validation frequency set to once every 30 epochs. Training will be terminated early if the loss on the validation set ceases to decrease continuously.

4 Numerical experiment

4.1 Data Acquisition and Preprocessing

This paper employs the terahertz time-domain spectroscopy system from Zomega Company. With a bandwidth ranging from 0.1 to 4 THz and a frequency resolution of 10 GHz, This system is employed to scan fabric samples point by point with a step size of 0.05 mm to obtain the terahertz time-domain spectral signals at each acquisition point are obtained through the schematic diagram in Fig. 3.

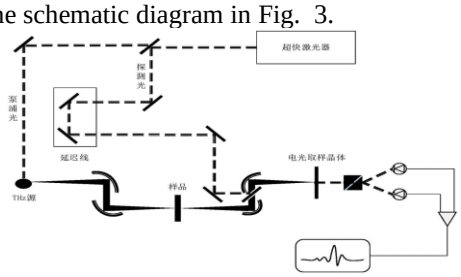


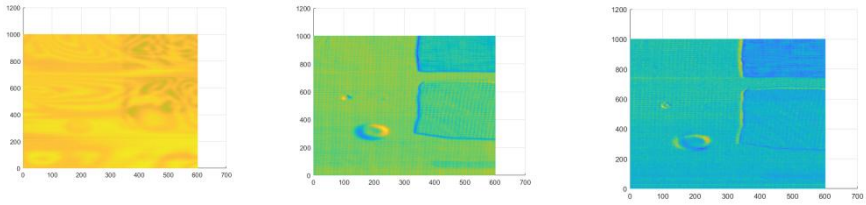
Fig. 3. Transmission terahertz time-domain spectroscopy system schematic diagram.

The original spectral data for eight different types of defects, the Savitzky-Golay filtering algorithm [23] was applied to reduce noise interference from the instrument and the environment, thereby enhancing the accuracy and reliability of the classification model. The basic equation of S-G filtering is as follows:

$$\hat{y}_i = \sum_{k=-m}^{m} c_k y_{i+k} , \tag{12}$$

where $\hat{y}_i$ is the smoothed value at i, y_{i+k} is the point of the original data, k is the offset within the window, c_k is the coefficient of the filter, and m is the half-width of the window.

A specific frequency imaging method was used to generate imaging results of the fabric samples at various frequencies. Images began to emerge at 0.4 THz, reached the highest clarity at 1.1 THz, and gradually blurred after 1.5 THz. Representative images are shown in Fig. 4.



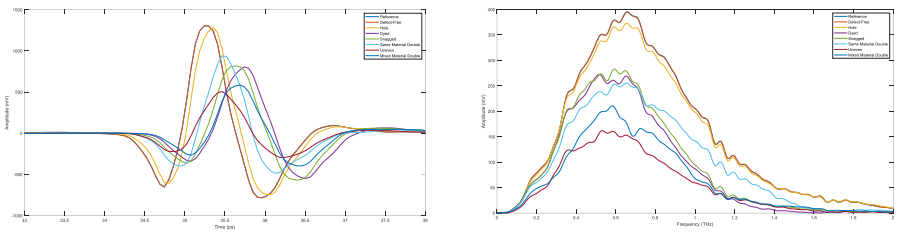
(a) (b) (c)

Fig. 4. Imaging results at different frequencies. (a)Imaging results at 0.4 THz,(b)Imaging results at 1.1 THz,(c)Imaging results at 1.5 THz.

The image with the highest clarity at 1.1 THz was selected as the basic data. Median filtering and grey stretching pre-processing were performed on this image to eliminate noise points and blurred regions and enhance defect characteristics. The processed image data and spectral data were paired one-to-one to form a dataset, which was then labelled. Specifically, the non-defective area was marked as 0; the polyester-cotton material overlay defect as 1; the cotton material double overlay defect as 2; stains as 3; holes as 4; hooks as 5; and twisted unevenness as 6. Given the similarity between holes and hooks, areas exhibiting features of both were defined as mixed defects and marked as 7. A total of 552,900 datasets, which were complete and contained multiple defect types, were divided into training, testing, and validation sets at a ratio of 7:1.5:1.5.

4.2 Experimental results and analyses

Using the data measured by the terahertz time-domain spectroscopy system, the time-domain and frequency-domain plots of the reference signal and the sample signal are depicted in Fig. 5 a. By comparing the time-domain plots of the signal passing through air (reference signal), defect-free fabrics, as well as fabrics with holes, dye stains, hooks, uneven twists, cotton double-layer overlays, and cotton-polyester laminates, a time delay and amplitude attenuation are observed between the time-domain reference signal and the sample signals. The time delay occurs because terahertz waves travel through samples with varying refractive indices, while the amplitude attenuation is caused by the absorption and random scattering of the samples. The corresponding frequency-domain plot in Fig. 5 b. clearly shows that the electric field of the sample signal is attenuated relative to the reference signal. This attenuation is attributed to the different absorption of terahertz waves by samples with various defects. These results demonstrate that different defects can be characterized by unique features within the terahertz spectral range, confirming the feasibility of terahertz time-domain spectroscopy imaging for fabric detection.

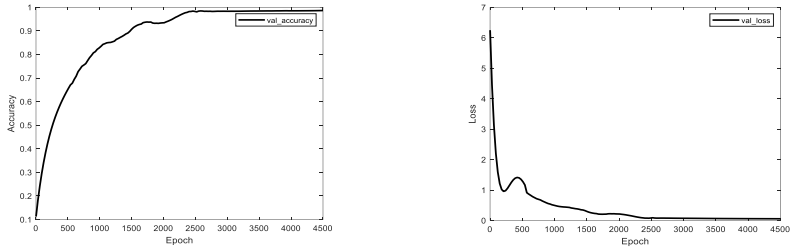


(a) (b)

Fig. 5. Time and frequency domain plots of the reference and sample signals. (a) Time-domain plots of reference and sample,(b) Frequency domain plots of reference and sample.

The terahertz time-domain spectroscopy system employed in the experiments had a nominal spectral range of 0.1-4.0 THz. To achieve a better signal-to-noise ratio, a spectral range of 0.1 to 1.5 THz was selected for this spectral analysis. The THz absorption coefficients of each defect were calculated using Eq. 11. The absorption coefficient data of different fabric defects were then fused with the image data and fed into the improved

CNN model. After several rounds of iterative training, as illustrated in Fig. 6 a., the accuracy continuously increased with the rise in the number of training rounds and stabilized at a high level. The improved CNN model achieved a classification accuracy of up to 99.4%. As shown in Fig. 6 b., the loss value decreased rapidly and gradually approached 0, indicating that the model possesses excellent learning ability and generalization performance and can accurately identify fabric defects. This effectively validates the research objective of efficiently and accurately classifying and identifying fabric defects.



(a) (b)

Fig. 6. Improved CNN model training accuracy versus loss function plot. (a)Accuracy graph,(b)Loss function plot.

4.3 Comparative Analysis of Models

To further validate the advantages of the improved CNN method proposed in this paper, a comparison with the traditional CNN model is conducted. The results indicate that although the traditional CNN model can utilize image information for classification, it fails to identify subtle and hidden defects. Its accuracy is 93.3%, approximately 6% lower than that of the improved CNN method.

To evaluate the model's classification performance more intuitively, this paper calculates the confusion matrices for the training, test, and validation sets of both the traditional and improved CNN models, as shown in **Fig. 7**. Specifically, coordinates 1 to 8 denote the classification of defect-free fabric, cotton-polyester material overlay defects, cotton material double overlay defects, stained defects, hole defects, hook defects, twisted and uneven defects, and mixed defects with both hole and hook features, respectively. **Fig. 7** a shows that the traditional CNN model achieves a classification accuracy of 93.5% on the training set, while **Fig. 7** b presents the improved CNN model's training-set confusion matrix with a classification accuracy of 99.3%. After fully training both models until their accuracies stabilize, they are applied to classify eight fabric categories, yielding the test-set confusion matrices in **Fig. 7** c and **Fig. 7** d. The traditional CNN model has a test accuracy of 93.3%, whereas the improved CNN model's test accuracy increases to 99.0%. Through constructing and analyzing the validation-set confusion matrices, the traditional CNN model has a validation accuracy of 93.4%, while the improved CNN model attains a perfect 100% validation accuracy, as depicted

in **Fig. 7** e and **Fig. 7** f. Evidently, the improved CNN model fully meets the classification and recognition requirements for the eight fabric defect types, with its accuracy significantly surpassing that of the traditional CNN model.

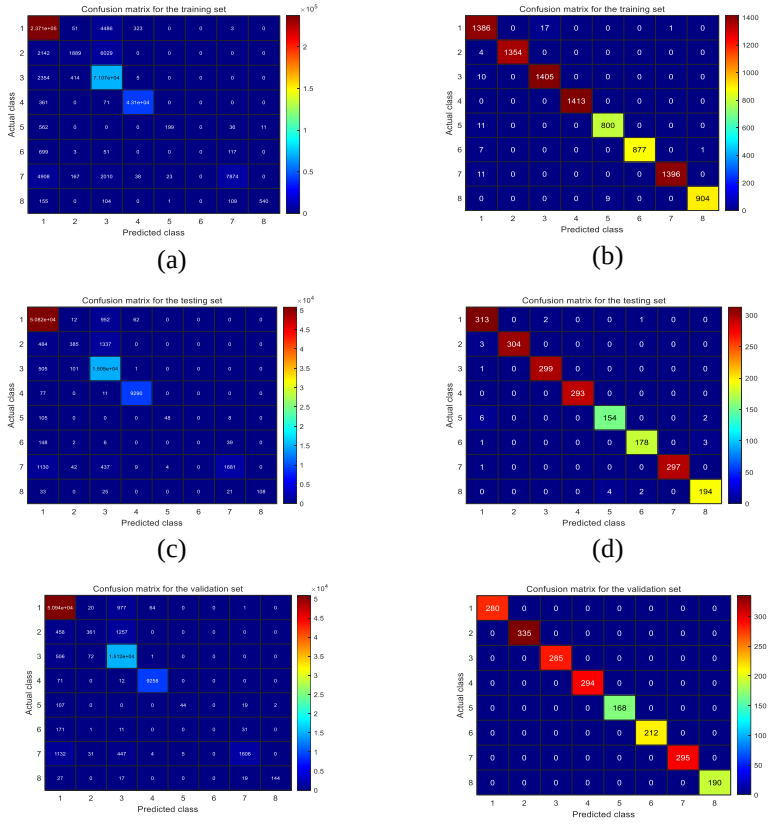


Fig. 7. Confusion matrices of traditional and improved CNN models. (a) (c) (e) are the confusion matrices for the training, validation, and test sets of the traditional CNN model, and (b) (d) (f) are the confusion matrices for the corresponding sets of the improved CNN model.

To conduct a more comprehensive evaluation of the classification model, this paper introduces the F1 score as an additional evaluation index, complementing the comparison of classification accuracy. As shown in Table 2, the improved CNN model outperforms the traditional CNN model in all four evaluation metrics: accuracy, precision, recall, and F1 score. This demonstrates that the improved CNN method can accurately differentiate various defect types and adapt to the diversity and complexity of fabric defects.

Table 2. Performance comparison of different classification models.

Classification model	Accuracy	Precision	Recall	F1-score
Traditional CNN	0.93	0.96	0.95	0.95
Improving CNN	0.99	0.99	0.99	0.99

5 Conclusion

In this paper, an improved convolutional neural network model based on multimodal data fusion is proposed to address the limitations of traditional convolutional neural networks in complex background defect recognition. Combining the advantages of terahertz time-domain spectral system, the feature extraction and classification performance is significantly improved by obtaining the deep fusion of absorption coefficients and image data, and combining the multilayer asymmetric convolutional kernel with jump connection design. Experimental results show that the improved convolutional neural network model achieves 99.4% accuracy and 0.99 F1 score in eight defect classification tasks, which is 6% and 4% higher than the traditional convolutional neural network method. It also performs well in terms of precision, recall, and confusion matrix accuracy. This study not only provides an efficient and accurate solution for defect recognition in complex scenes, but also has high application value and research significance in the future.

Acknowledgments. This study was supported by the Key Technology and Application of Government Large-scale Modelling Project (2024CXGC010111) and Jinan Municipal School Integration and Development Strategy Project 'Hongmeng +' Innovative and Applied Digital Talent Cultivation Characteristic College (JNSX2024043).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Zahra, A., Amin, M., El-Samie, F.E.A., et al.: Efficient utilization of deep learning for the detection of fabric defects. Neural Computing and Applications 36(11), 6037-6050(2024)
2. Kayumov, A., Sobirov, M., Musayev, K.: Methods of fabric defect detection using expert systems-a systematic literature review. In: E3S Web of Conferences. EDP Sciences, 538, 04015 (2024)
3. Rahman, M.S.U., Hassan, O.S., Mustapha, A.A., et al.: Inspection of thick composites: a comparative study between microwaves, X-ray computed tomography and ultrasonic testing. Nondestructive Testing and Evaluation 39(7), 2054–2071 (2024)
4. Miqoi, N., Pomarede, P., Meraghni, F., et al.: Detection and evaluation of barely visible impact damage in woven glass fabric reinforced polyamide 6.6/6 composite using ultrasonic imaging, X-ray tomography and optical profilometry. International Journal of Damage Mechanics 30(3), 323–348 (2021)

5. Chen, M., Jiang, X., Huang, J., et al.: Flexible wearable Ti3C2Tx composite carbon fabric textile with infrared stealth and electromagnetic interference shielding effect. Advanced Optical Materials 12(4), 2301694 (2024)
6. Wang, H., Luo, H., Zhang, X., et al.: Automatic defect detection of carbon fiber woven fabrics using machine vision. Mechanics of Advanced Materials and Structures 31(28), 10921–10934 (2024)
7. Davies, A.G., Linfield, E.H., Johnston, M.B.: The development of terahertz sources and their applications. Physics in Medicine & Biology 47(21), 3679 (2002)
8. Prokscha, A., Sheikh, F., Kolpatzeck, K., et al.: Terahertz Insights Into Fabric Look-Through. In: 2024 15th German Microwave Conference. IEEE, pp. 241–244 (2024)
9. Xiang, A., Zhang, J., Yang, Q., et al.: Research on splicing image detection algorithms based on natural image statistical characteristics. arXiv preprint arXiv:2404.16296 (2024)
10. Jeyaraj, P.R., Samuel Nadar, E.R.: Computer vision for automatic detection and classification of fabric defect employing deep learning algorithm. International Journal of Clothing Science and Technology 31(4), 510–521 (2019)
11. Chauhan, R., Ghanshala, K.K., Joshi, R.C.: Convolutional neural network (CNN) for image detection and recognition. In: 2018 first international conference on secure cyber computing and communication (ICSCCC). IEEE, pp. 278–282 (2018)
12. Wang, X., Xu, Y., Cui, Y., et al.: DA-CNN-based similar terahertz signal identification for intelligent characterization of internal debonding defects of composites under high-resolution mode. Composite Structures 322, 117412 (2023)
13. Kim, H.S., Park, D.W., Kim, S.I., et al.: Non-destructive detection of thin micro-defects in glass reinforced polymer composites using a terahertz electro–magnetic wave based on a convolution neural network. Composites Part B: Engineering 257, 110694 (2023)
14. Zhang, H., Wu, Y., Lu, S., et al.: A mixed-attention-based multi-scale autoencoder algorithm for fabric defect detection. Coloration Technology 140(3), 451–466 (2024)
15. Mengiste, E., Mannem, K.R., Prieto, S.A., et al.: Transfer-learning and texture features for recognition of the conditions of construction materials with small data sets. Journal of Computing in Civil Engineering 38(1), 04023036 (2024)
16. Dorney, T.D., Baraniuk, R.G., Mittleman, D.M.: Material parameter estimation with terahertz time-domain spectroscopy. J. Opt. Soc. Am. A Opt. Image Sci. Vis. 18, 1562–1571 (2001)
17. Duvillaret, L., Garet, F., Coutaz, J.L.: A reliable method for extraction of material parameters in terahertz time-domain spectroscopy. IEEE J. Sel. Top. Quantum Electron. 2, 739–746 (2002)
18. Bassiri, S., Papas, C.H., Engheta, N.: Electromagnetic wave propagation through a dielectric-chiral interface and through a chiral slab. J. Opt. Soc. Am. A 5, 1450–1459 (1988)
19. Zhang, Y., Peng, X.H., Chen, Y., Chen, J., Curioni, A., Andreoni, W., Nayak, S.K., Zhang, X.C.: A first principle study of terahertz (THz) spectra of acephate. Chem. Phys. Lett. 452, 59–66 (2008)
20. Nsengiyumva, W., Zhong, S., Zheng, L., et al.: Theoretical and experimental analysis of the dielectric properties of 3D orthogonal woven GFRP composites in the terahertz frequency range. Optik 260, 169105 (2022)
21. Liu, D., Lu, T., Qi, F.: A reliable method for removing Fabry–Perot effect in material characterization with terahertz time-domain spectroscopy. IEEE Transactions on Terahertz Science and Technology 10(5), 443–452 (2020)
22. Qi-feng, H., Jian, C.: Research of terahertz time-domain spectral identification based on deep learning. Spectroscopy and Spectral Analysis 41(1), 94–99 (2021)

23. Qiu, W., Wang, X., Ma, D., et al.: A recurrent neural network for adaptive filtering in terahertz time-domain spectroscopy. Infrared Physics & Technology 138, 105256 (2024)