Classification of Infrasound Events Based on Multilevel Wavelet Transforms

Jiang Yue and Cui YiDong

September 15, 2023

Classification of Infrasound Events based on Multilevel Wavelet Transforms

Yue Jiang,Yidong Cui

Network and Communication Software Center

Beijing University of Posts and Telecommunications, Beijing 100876, China

jiangyuebupt@bupt.edu.cn, cyd@bupt.edu.cn

***Abstract*—Classification and recognition methods for infrasound events are widely used in various fields. Although traditional classification methods have made attempts to handle infrasound signals, there are challenges in dealing with skewed training samples in the classification model and achieving accurate classification for events with limited samples due to the rarity of certain infrasound events. To address the classification problem with a small number of chemical explosion samples in the training set, this paper proposes an improved deep convolutional neural network model for the classification of infrasound signals. In the comparative analysis, we compare the improved deep convolutional neural network with the standard LeNet and ResNet models. The experimental results demonstrate that the proposed classification model achieves similar performance to the advanced ResNet model in terms of test recognition rate, while requiring fewer covariates. This provides an advantage over previous algorithms.**

Keywords—* *infrasound event classification, convolutional neural network, sample skewness, multilevel wavelet transform

I. Introduction

Infrasound is a low-frequency sound wave with energy below the threshold of human hearing, with a frequency of about 20 Hz. Infrasound is generated by a variety of natural and anthropogenic sources, and, because of its long wavelength and low frequency, it propagates on land, in the oceans and in the atmosphere, as compared with high-frequency sound waves. Many physical phenomena are accompanied by low-frequency infrasound signals during their occurrence and development, such as earthquakes, typhoons, lightning, volcanic eruptions, and tsunamis in natural activities, as well as nuclear and chemical explosions, rocket launches, and airplane takeoffs in human activities [1]. Due to these sources and propagation characteristics of infrasound, infrasound is a key technology for the detection of natural disasters, nuclear and chemical explosions, and other events generated by anthropogenic or natural sources on a regional and global scale.

Since the 1950s, scholars at home and abroad have explored the task of classifying infrasound events in depth, mainly utilizing acoustic signal processing theory to extract features in the time and frequency domains, respectively, which are then used for classification and recognition. Traditional machine learning methods are well suited for the recognition of nonlinear patterns of infrasound events [4], but it is difficult to obtain a large amount of data for some rare infrasound events, such as chemical explosions, etc., and thus the infrasound event recognition dataset suffers from sample skewing, which leads to the poor effect of training the generalized model to classify the categories that are missing training samples. Some studies have pointed out that the wavelet transform is one of the ways to obtain the infrasound event "fingerprint", and experiments have proved that the use of a specific wavelet basis function and wavelet packet decomposition layer number, for some categories of infrasound signals can be divided into better. At the same time, the convolutional neural network has a powerful feature extraction and learning ability, can be extracted to the artificial features can not be described, and complete the classification and recognition, in the digital recognition, image classification and recognition, noise removal and other fields have been well used. We first consider using a wavelet convolutional layer to replace the first convolutional layer of a standard CNN, and compare the effects of using different wavelet basis functions and different network structures on the classification results. Experiments demonstrate that using a wavelet convolutional layer to replace the first convolutional layer of a standard CNN improves the classification accuracy when there are sufficient samples, but the classification accuracy and checking completeness are still low for small-sample categories. We observe that multilevel wavelet packet decomposition can extract more time-frequency domain information, and higher order wavelet packet decomposition can be better separable for certain types of infrasound events. (The method of using one wavelet convolution layer as the first convolution layer of the CNN is similar to the process of traditional classification methods and can be regarded as a preprocessing process for infrasound signals.) Therefore, we used multilevel wavelet packet decomposition to extract features in parallel with the CNN; specifically, we modified the backbone of the standard CNN using multilevel wavelet packet decomposition to enable the incorporation of wavelet packet component features from all levels during the convolution process. Finally, we conducted several experiments to demonstrate that the modified CNN model using multilevel wavelet decomposition achieves better results in the problem of infrasound event classification in sample skewed scenarios.

II. Related Work

A. Methods for Infrasound Event Classification

The majority of current algorithms for infrasound event recognition employ simple models such as Support Vector Machine (SVM), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs). With the emergence of artificial neural networks, there has been a gradual shift towards combining complex features with artificial neural networks for the classification and recognition of infrasound events, enabling more effective improvement in infrasound event recognition. Albert et al.[1]conducted experimental analyses and demonstrated that the standard CNN achieves comparable classification accuracy to the SVM algorithm. However, due to its simplicity and physical interpretability, with features correlated to the physical properties of waveforms, the SVM is deemed the preferred method for their dataset as described in their study. Bishop et al. [4]developed a deep learning-based method for infrasound detection and classification, employing a CNN with a self-attentive layer to identify fixed and non-fixed signals in the results of infrasound array processing. The improved model yields more reliable detection results for infrasound signals compared to raw waveform data. Another physically-based approach, introduced by Alex Witsil et al. [18], generated 28,000 synthetic events from realistic atmospherically propagating infrasound sources for training Machine Learning (ML) classifiers. The experiments showed that training exclusively on a synthetic physics-based dataset enhanced model performance in detecting explosions and improved ML performance in classifying domains with limited training data. In this paper, Tan Xiaofeng et al. utilized the Welch power spectral transform for feature extraction, followed by the use of an improved CNN for classification. Their experimentation involved a dataset comprising 815 chemical explosions and natural seismic infrasound signals (referred to as earthquakes), analyzing the classification performance using Backpropagation (BP) networks and a one-dimensional LeNet-5 network. .The experimental results demonstrate that the improved CNN outperforms the aforementioned methods。.

B. *Feature extraction*

Wavelet analysis is a time-frequency signal analysis technique that differs from traditional signal analysis methods such as the Fourier transform. The Fourier transform provides a global view of signal characteristics in either the time or frequency domain, but fails to capture the local time-frequency characteristics of non-linear and non-smooth signals, which are essential and unique features of such signals.

In the example of a 3-layer wavelet transform shown in Fig 1, the signal can be decomposed into low-frequency (approximation coefficient, denoted as An) and high-frequency (detail coefficient, denoted as Dn) components at different levels or scales.

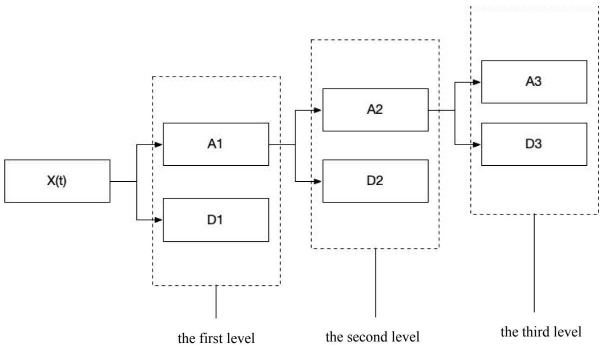

Fig. 1. Multi-layer wavelet decomposition

Suppose $X(t)$ is the signal to be transformed, the continuous wavelet transform can be expressed as, the

$$W(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} X(t)\varphi^{*}\left(\frac{t-b}{a}\right)dt \qquad (1)$$

However, due to the infinite length of continuous wavelet functions, the obtained coefficients are also infinitely long, making it impractical for computations. To address this, the discrete wavelet transform was developed, using wavelet functions of finite length (e.g., Daubechies series) to discretize the signal and decompose it through discrete convolution. The discrete wavelet transform possesses the multi-resolution characteristic, allowing the decomposition of a signal into different scales, making it widely applicable in various signal processing tasks such as compression, denoising, and feature extraction.

Because wavelet analysis has good time-frequency analysis ability, it is widely used in the feature extraction of non-smooth, nonlinear signals and signal processing in target recognition, which provides ideas for infrasound signal processing.Chilo et al [6] and others pointed out that the wavelet transform, if correctly combined with subsequent post-processing and recognition tools, is one of the ways to obtain the "fingerprint" of infrasound events. "Chilo et al. Jiang Nan et al [7] proposed the use of wavelet packet decomposition method of infrasound signal feature extraction research, and pointed out that the selection of a specific wavelet basis function and wavelet packet component ratio features on the nuclear and chemical explosions infrasound signals can be divided into better.

III. METHODOLOGY

The Db4 wavelet basis function, a classical Daubechies wavelet basis function, possesses several properties that make it suitable for signal decomposition of infrasound signals. Firstly, it has a compact support interval in the time domain, allowing it to effectively capture rapid changes in these signals. Infrasound signals typically exhibit narrow spectral bandwidths and fast variations, making a compactly supported wavelet function ideal for extracting such features. Secondly, the Db4 wavelet basis function concentrates a significant amount of energy in the high-frequency band, enabling it to extract the high-frequency components of infrasound signals more accurately. These signals often contain subtle high-frequency detail information, which can be decomposed and reconstructed with greater precision using the Db4 wavelet basis function. Additionally, the Db4 wavelet basis functions form an orthogonal set of basis functions, ensuring independence and mutual uncorrelation between the decomposition coefficients. This orthogonality enables better representation and analysis of infrasound signals.

The LeNet-5 network, proposed by Lecun in 1998, is a classic convolutional neural network (CNN) well-suited for handwritten character recognition. It comprises seven layers (excluding inputs), each containing trainable parameters (weights). The network includes two sets of convolutional and pooling layers for feature extraction, and a fully-connected layer for mapping learned features to sample labelling space, acting as a "classifier". The LeNet-5 network is relatively simple, with a small number of convolutional layers and parameters, resulting in faster learning.In practical applications, we employ the Db4 wavelet basis function as the basis for wavelet decomposition to decompose input infrasound signals at different levels. The resulting coefficients are incorporated into the structure of the LeNet convolutional neural network to enhance feature extraction and classification of infrasound signals.

We propose two methods to improve the LeNet structure: LeNet networks enhanced using single-level wavelet

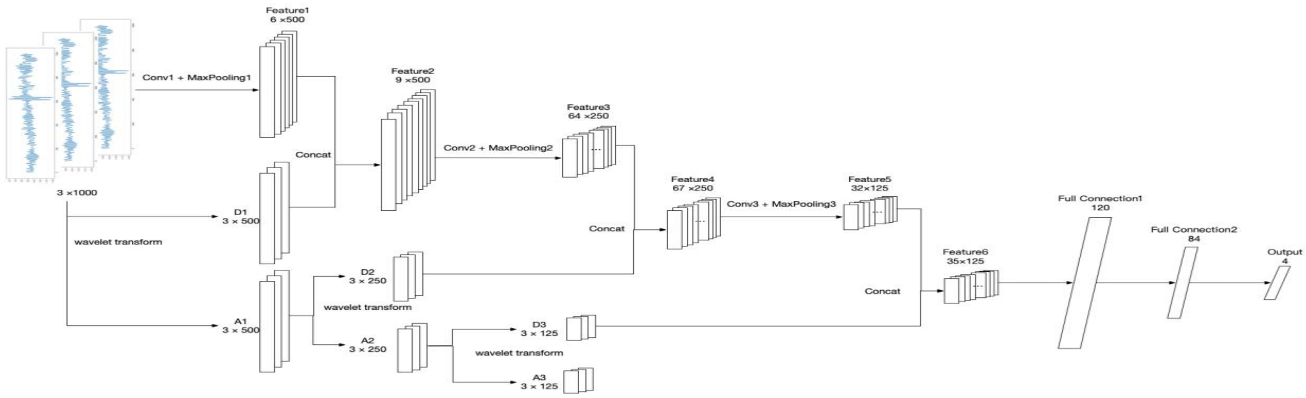

Fig. 2. Improved LeNet model with multi-level wavelet Transforms

decomposition and LeNet networks enhanced using multi-level wavelet decomposition. The subsequent sections will provide detailed explanations of these two methods.

A. *Enhancing LeNet Structure with Single-Level Wavelet Decomposition*

Due to the one-dimensional nature of infrasound signals, a one-dimensional convolution kernel is uniformly employed in the LeNet-5 network to facilitate infrasound data processing. Furthermore, the LeNet structure is enhanced by incorporating a single-level wavelet decomposition. Specifically, the convolution operation in the initial layer of LeNet is replaced with wavelet decomposition, implemented through a single-stage wavelet decomposition. Infrasound arrays typically consist of a ternary array and a quintuple array, which find applications in tasks such as sound source localization and classification. In practical scenarios, processing steps such as correction and denoising are usually performed on infrasound arrays to mitigate errors and enhance measurement accuracy. Taking the input of a time-series from an infrasound 3-element array, each of the three input sequences undergoes a wavelet decomposition, resulting in three sets of low and high frequency subbands. The three low-frequency subbands are then combined and used as the input for the subsequent LeNet layer. After two convolution and pooling operations, the feature map is obtained. The high-frequency subbands are fused with the feature map, which is then input to the fully connected layer. The sigmoid activation function is applied in this layer, while the softmax activation function is utilized in the output layer to produce the final classification results.

The specific model structure is as follows:

- Input layer: 3 channels of raw infrasound signal
- Layer 1: The input image is subjected to one wavelet decomposition to obtain two subbands, low frequency and high frequency. The low-frequency sub-band is taken as input and the feature map is obtained by convolution operation.
- Layer 2: Pooling layer that downsamples the feature maps of the first layer.
- Layer 3: The output of layer 2 is taken as input and a new feature map is obtained after convolution operation.
- Layer 4: Pooling layer, down-sampling the feature maps from layer 3.
- Layer 5: Weighted fusion of the Layer 4 output and the HF subbands of Layer 1.
- Layer 6: Fully connected layer, the output of layer 6 is multiplied with the weight matrix and a bias term is added to get the new feature vector.
- Layer 7: Fully connected layer, the output of layer 7 is again multiplied with the weight matrix and a bias term is added to get the final classification result.

In fusing the HF subbands and feature maps, weighted averaging is used, i.e., the HF subbands and feature maps are multiplied by different weighting coefficients and then summed to obtain the new inputs. The weighting coefficients were determined experimentally for different types of infrasound events.

B. *Enhancing LeNet Structure Using Multilevel Wavelet Decomposition*

In practical single-stage wavelet decomposition to improve the LeNet structure, the signal is decomposed into a set of low-frequency subbands and a set of high-frequency subbands. Among them, the low-frequency subbands contain most of the energy of the signal and the high-frequency subbands contain the high-frequency features of the signal. By extracting features from the low-frequency sub-band coiler and fusing the high-frequency sub-bands for analysis and processing, infrasound events can be classified. However, since the single-stage wavelet packet decomposition is performed only once, the number of high-frequency subbands obtained is small and may not adequately describe the high-frequency features in the infrasound signal. Multi-level wavelet packet decomposition is a multi-level division of frequency bands, which, for infrasound signals, can decompose the low-frequency portion of infrasound signals and the high-frequency portion that is not subdivided in the multiresolution analysis, and is able to adaptively select the corresponding frequency bands according to the characteristics of the infrasound signals to improve the time-frequency resolution. Because of the different spectral distribution characteristics and source mechanisms of various infrasound signals, wavelet packet decomposition has a wide range of applications for infrasound signals. In contrast, multilevel wavelet packet decomposition can further decompose the signal into more subbands of scales and frequencies, thus more comprehensively describing the features in the infrasound signals, including high-frequency features.

The LeNet model is improved using multilevel wavelet decomposition as shown in Fig.2 The specific model structure is as follows:

- Input layer: 3 channels of raw infrasound signal
- Layer 1: The input image is subjected to the first wavelet decomposition to obtain two subbands, low-frequency A1 and high-frequency D1. The high-frequency subband D1 and the 3-channel original infrasound signal are used as inputs, and the feature map is obtained by convolution operation.
- Layer 2: Pooling layer that downsamples the feature maps of the first layer.
- Layer 3: The low-frequency subband A1 obtained from the first wavelet decomposition is subjected to the second wavelet decomposition to obtain two subbands, low-frequency A2 and high-frequency D2, and the high-frequency subband D2 and the feature map outputted from the second layer are used as inputs, and new feature maps are obtained after convolution operation.
- Layer 4: Pooling layer, down-sampling the feature maps from layer 3.
- Layer 6: Expand the output of layer 4 into a one-dimensional vector.
- Layer 7: Fully connected layer, the low-frequency subband A2 obtained from the third wavelet decomposition is subjected to the third wavelet decomposition to obtain two subbands, low-frequency A3 and high-frequency D3, and the outputs of the high-frequency subbands D3 and the sixth layer are multiplied by the weight matrix and the bias term is added to obtain the new eigenvectors.
- Layer 8: Fully connected layer, where the output of layer 7 is again multiplied with the weight matrix and a bias term is added to get the final classification result.

Compared with the "LeNet-5" network, this paper makes the following improvements to the LeNet-5 network:

The first convolutional layer uses a wide convolutional kernel, the widened convolutional kernel allows for a larger sensory field and the wide kernel in the first convolutional layer suppresses high frequency noise. So, in the improved LeNet-5 network, the convolution kernel in the first layer is changed to 1×500 . Subsequent convolutional layers use more smaller convolutional kernels to extract more detailed features.

Multi-level wavelet classification provides higher resolution and more detailed frequency analysis, using multi-level wavelet decomposition to improve the LeNet-5 network, each level of wavelet decomposition decomposes the signal into an approximate component (low-frequency portion) and a detailed component (high-frequency portion), combining the high-frequency portion with the input of that level of convolution and extracting the features through the convolution, and then the low-frequency portion is further decomposed at the next level to extract more detailed information, and the above operation is repeated for the low-frequency part and the high-frequency part of the decomposition output. This level-by-level decomposition provides a more comprehensive signal analysis.

IV. Experiment

A. Data pre-processing

The infrasound waveform dataset in this paper is collected from multiple 3-element arrays of four main categories of infrasound events, namely typhoons, lightning, chemical explosions, and microbarometric pressure. The infrasound signal data is sampled at 100 hertz (Hz), i.e., 100 samples per second. The types of infrasound events and the amount of data included in the infrasound event identification dataset are shown in Table 3-1. Each event is collected by 3 devices, so each event has 3 infrasound signals, and the number of infrasound signals in each category is shown in Table 3-1. The original data set is txt text data. The txt text contains infrasound signal data and the timestamp information of each sampling point.

TABLE I. Statistics on the original data set

Type of incident	stage presence, poise	thunder and lightning	chemical explosion	microbarometer
Number of events	12	87	5	22
Number of infrasound signals	36	261	15	66

Acquired infrasound signals may receive electronic thermal noise from the acquisition device, non-linear response of the transducer, environmental noise, and other disturbances during the signal acquisition process, which may result in unwanted frequency components or interference in the signal. Filtering can help to remove or attenuate these interferences so that the signal is clearer and easier to interpret and analyse. Infrasound is a 0-20 Hz sound wave, so in this paper we use low-pass filtering to separate the target signal from the noise and improve the signal-to-noise ratio of the infrasound signal.

Due to the long duration of typhoon and micropressure events, the infrasound signal data for each typhoon and micropressure event in the original infrasound event identification dataset are sampled over a long period of time, i.e., the number of sampling points contained in each event is high. For each infrasound signal data, every 500 sampling points are taken as a new data, and those less than 500 sampling points are directly discarded. The total number of events processed by this method is 9640.

If the predicted value y and the true value y_ are the same, then the network is considered to have correctly identified the infrasound event. By comparing the predicted value y and the true value y_ of the neural network output, four statistics can be obtained: Ture Positive (TP, true value 1 and predicted by the network to be 1), False Positive (FP, true value 0 but predicted by the network to be 1), False Negative (FN, true value 1 but predicted by the network to be 0), Ture Negative (TN, true value 0 and predicted as 0 by the network). Based on the four statistics of TP, FP, FN, and TN, four evaluation metrics can be obtained, namely, Accuracy, Precision, Recall, and F1-Score. Using the above four indicators can provide a more reliable analysis of network accuracy and stability based on the experimental results.

F1-Score is a metric used to evaluate the performance of classification models, which combines the accuracy and recall of a model to provide a comprehensive assessment of model performance. Its calculation formula is

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \qquad (2)$$

The value of F1-Score ranges from 0 to 1, with closer to 1 indicating the model's ability to classify positive instances.F1-Score combines accuracy and recall, and therefore provides a more comprehensive assessment of model performance for certain application scenarios. In the case of infrasound event classification, a sample-skewed scenario, there is a category imbalance in the dataset being processed, and chemical explosions are relatively few in number compared to other events, so if the classifier tends to predict the majority of the categories, the accuracy may be high, but the recall will be low. At the same time, the problem of classifying events of all types in infrasound requires both accuracy and coverage of the prediction results, and both accuracy and recall are considered important metrics.F1-score will consider both accuracy and recall, and therefore can better evaluate the performance of the model on unbalanced datasets.

B. Experiments and Analysis of Results

In order to compare the effectiveness of the Improved LeNet model with multi-level wavelet decomposition and the Improved LeNet model with single-level wavelet decomposition on the infrasound event classification problem, we designed a series of comparison experiments. At the same time, we reduced the dataset and observed the performance of the models with reduced data volume to evaluate the classification effectiveness of the improved LeNet model. In addition, we compare the two models with the state-of-the-art ReNet model.

The ResNet model has obvious advantages in terms of network depth, etc., but due to its demand for a large amount of training data, the complexity of model training, and the large number of parameters, its training conditions are more demanding.The parameter comparison between the ResNet model and the LeNet model is shown in the TABLE II below, and the number of parameters in the ResNet model is about 22 times as many as that of the LeNet and its improved model.

TABLE II. COMPARISON OF THE NUMBER OF MODEL PARAMETERS

modelling	LeNet	Single-WT-LeNet	Multi-level-WT-LeNet	ResNet-50
quantity of participants	11,586	11,586	11,586	Approx. 256,000

C. Comparative experiments with different models

The training loss, training and testing accuracies of the four types of models are shown in Fig.3, which show that the training loss decreases slowly but converges slowly in the normal LeNet model. The single-stage wavelet transform-improved LeNet model shows faster convergence during training, and the loss decreases rapidly. Meanwhile, the multi-level wavelet transform improved LeNet model performs better in terms of training loss because it has higher representation capability and better feature extraction. In contrast, the normal ResNet model shows the best convergence performance with very fast training loss reduction and the final loss convergence result is similar to the multilevel wavelet transform improved LeNet.

The regular LeNet model has similar accuracy to the other models on the training set, but performs poorly on the validation set. The single-level wavelet transform improved LeNet model showed better performance in terms of test accuracy, achieving about 98% accuracy. The multi-stage wavelet transform improved LeNet model further improved the accuracy compared to the single-stage wavelet transform, achieving a similar accuracy to ResNet in the testing phase, close to 100%. The regular ResNet model showed the best performance in both training and testing accuracy, with a stable accuracy close to 100%.

Taken together, the single-level wavelet transform-improved LeNet improves the training and testing accuracy relative to the regular LeNet, but still lags behind the multi-level wavelet transform-improved LeNet and ResNet models. The multilevel wavelet transform-improved LeNet achieves similar performance to ResNet, indicating that the wavelet transform has better results for image classification tasks. Whereas, the regular ResNet model performs the best among these models with the fastest convergence speed and the highest training and testing accuracies.

However, for chemical explosion events, the total number of infrasound signals is 15 in total, and its proportion in the total samples is the smallest, so the precision rate, recall rate and F1-Score value of each model for the classification of chemical explosion events are discussed separately in order to compare the classification performance of the models. From Table 3, it can be seen that the LeNet model improved by multilevel wavelet transform and the ReNet model perform similarly, and both can extract more features in small-sample events and show better classification results.

1) Comparative experiments with reduced datasets

In the same experimental setting, we trained and recorded the loss values as well as the training and validation accuracies by reducing the dataset to 50% of the original size (i.e., containing only 4820 samples). As shown in Figs. 7~9, we can observe the following: after reducing the number of samples, the improved LeNet model with single-stage wavelet decomposition is similar to the regular LeNet model in terms of convergence speed, training and validation accuracies. However, it is worth noting that the multi-level wavelet decomposition improved LeNet model still has a significant advantage in classification accuracy. Its validation set accuracy reaches a level of about 99% with close to 100% accuracy in the training set.

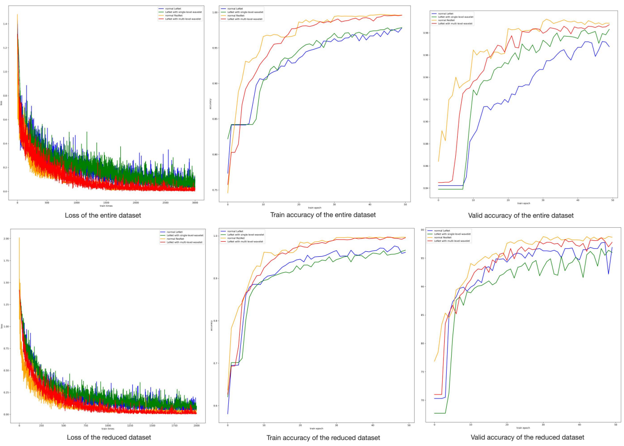

Fig. 3. Training loss and accuracy of the model

V. CONCLUSION

We propose an improved LeNet model for the small-sample event classification problem by combining different levels of wavelet decomposition with an improved LeNet network structure. To evaluate its performance, we used the LeNet model, the improved LeNet model, and the traditional ResNet model for comparison in our experiments. The experimental results show that the improved LeNet model using multilevel wavelet decomposition achieves a similar level to the ResNet model in terms of validation set accuracy, precision and recall for small-sample event classification. It is worth noting that the number of parameters in this improved LeNet model is only 1/22 compared with the ResNet model.It is demonstrated experimentally that our proposed improved LeNet model using multilevel wavelet decomposition can achieve better small-sample event classification with a smaller number of parameters. This research result provides new methods and ideas for solving the small-sample event classification problem.

There are multiple types of infrasound sources in nature, and infrasound events are not only limited to the four cases in the current training dataset. Deep learning, as an effective method for dealing with multi-classification problems, can be used to achieve the classification and identification of multiple types of infrasound events. To achieve this goal, a database containing different types of infrasound events (e.g., nuclear explosions, natural earthquakes, chemical explosions, mine explosions, rocket launches, etc.) is required. In the next step of our research, we will collect data from more types of infrasound events and study the feature extraction method for each type whose separability is the best in order to complete the work of classifying and recognising multiple types of infrasound events.

REFERENCES

[1] S. Albert and L. Linville, ‘Benchmarking Current and Emerging Approaches to Infrasound Signal Classification’, Seismological Research Letters, vol. 91, no. 2A, pp. 921–929, Mar. 2020, doi: 10.1785/0220190116.

[2] S. Arrowsmith, G. Euler, O. Marcillo, P. Blom, R. Whitaker, and G. Randall, ‘Development of a robust and automated infrasound event catalogue using the International Monitoring System’, Geophysical Journal International, vol. 200, no. 3, pp. 1411–1422, Jan. 2015, doi: 10.1093/gji/ggu486.

[3] J. D. Assink, G. Averbuch, P. S. M. Smets, and L. G. Evers, ‘On the infrasound detected from the 2013 and 2016 DPRK’s underground nuclear tests’, Geophys. Res. Lett., vol. 43, no. 7, pp. 3526–3533, Apr. 2016, doi: 10.1002/2016GL068497.

[4] J. W. Bishop, P. S. Blom, J. Webster, W. Reichard-Flynn, and Y. Lin, ‘Deep learning categorization of infrasound array data’, The Journal of the Acoustical Society of America, vol. 152, no. 4, pp. 2434–2445, Oct. 2022, doi: 10.1121/10.0014903.

[5] T. Boczar, D. Zmarzły, M. Kozioł, and D. Wotzka, ‘Application of Correlation Analysis for Assessment of Infrasound Signals Emission by Wind Turbines’, Sensors, vol. 20, no. 23, p. 6891, Dec. 2020, doi: 10.3390/s20236891.

[6] N. Brachet, D. Brown, R. Le Bras, Y. Cansi, P. Mialle, and J. Coyne, ‘Monitoring the Earth’s Atmosphere with the Global IMS Infrasound Network’, in Infrasound Monitoring for Atmospheric Studies, A. Le Pichon, E. Blanc, and A. Hauchecorne, Eds., Dordrecht: Springer Netherlands, 2010, pp. 77–118. doi: 10.1007/978-1-4020-9508-5_3.

[7] J. Chilo et al., ‘Filtering and extracting features from infrasound data’, in 14th IEEE-NPSS Real Time Conference, 2005., Stockholm, Sweden: IEEE, 2005, p. 5 pp. doi: 10.1109/RTC.2005.1547494.

[8]

[9] Z. A. Khan, T. Hussain, F. U. M. Ullah, S. K. Gupta, M. Y. Lee, and S. W. Baik, ‘Randomly Initialized CNN with Densely Connected Stacked Autoencoder for Efficient Fire Detection’, Engineering Applications of Artificial Intelligence, vol. 116, p. 105403, Nov. 2022, doi: 10.1016/j.engappai.2022.105403.

[10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, ‘Gradient-based learning applied to document recognition’, Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[11] T. Li et al., ‘WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis’, IEEE Trans. Syst. Man Cybern, Syst., vol. 52, no. 4, pp. 2302–2312, Apr. 2022, doi: 10.1109/TSMC.2020.3048950.

[12] J. Marty, ‘The IMS Infrasound Network: Current Status and Technological Developments’, Infrasound Monitoring for Atmospheric Studies: Challenges in Middle Atmosphere Dynamics and Societal Benefits. Springer International Publishing, Cham, pp. 3–62, 2019. [Online]. Available: https://doi.org/10.1007/978-3-319-75140-5_1

[13] T. Perol, M. Gharbi, and M. Denolle, ‘Convolutional neural network for earthquake detection and location’, Sci. Adv., vol. 4, no. 2, p. e1700578, Feb. 2018, doi: 10.1126/sciadv.1700578.

[14] J. Redmon and A. Farhadi, ‘YOLOv3: An Incremental Improvement’. arXiv, Apr. 08, 2018. Accessed: May 24, 2023. [Online]. Available: http://arxiv.org/abs/1804.02767

[15] X. Song, Y. Cong, Y. Song, Y. Chen, and P. Liang, ‘A bearing fault diagnosis model based on CNN with wide convolution kernels’, J Ambient Intell Human Comput, vol. 13, no. 8, pp. 4041–4056, Aug. 2022, doi: 10.1007/s12652-021-03177-x.

[16] L. Toney, D. Fee, A. Witsil, and R. S. Matoza, ‘Waveform Features Strongly Control Subcrater Classification Performance for a Large, Labeled Volcano Infrasound Dataset’, The Seismic Record, vol. 2, no. 3, pp. 167–175, Jul. 2022, doi: 10.1785/0320220019.

[17] A. Witsil, D. Fee, J. Dickey, R. Peña, R. Waxler, and P. Blom, ‘Detecting Large Explosions With Machine Learning Models Trained on Synthetic Infrasound Data’, Geophysical Research Letters, vol. 49, no. 11, Jun. 2022, doi: 10.1029/2022GL097785.

[18] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, ‘A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals’, Sensors, vol. 17, no. 2, p. 425, Feb. 2017, doi: 10.3390/s17020425.

[19] H. Zhou, W. Dong, Y. Liu, and J. Chen, ‘Breaking Through the Haze: An Advanced Non-Homogeneous Dehazing Method based on Fast Fourier Convolution and ConvNeXt’. arXiv, May 07, 2023. Accessed: Jun. 18, 2023. [Online]. Available: http://arxiv.org/abs/2305.04430