Towards an Interpretable CNN Model for the Classification of Lightning Produced VLF/LF Signals

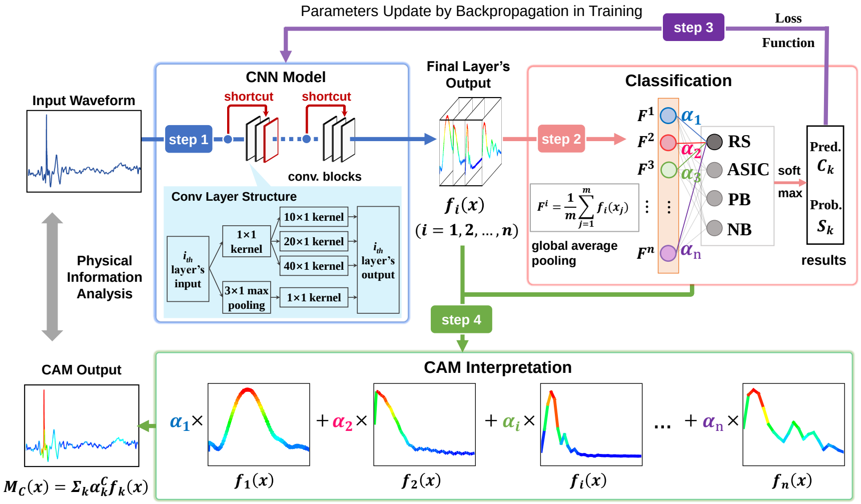
Lilang Xiao[1], Weijiang Chen[2], Yu Wang[3], Kai Bian[4], Zhong Fu[5], Nianwen Xiang[6], Hengxin He[7], and Yang Cheng[5]

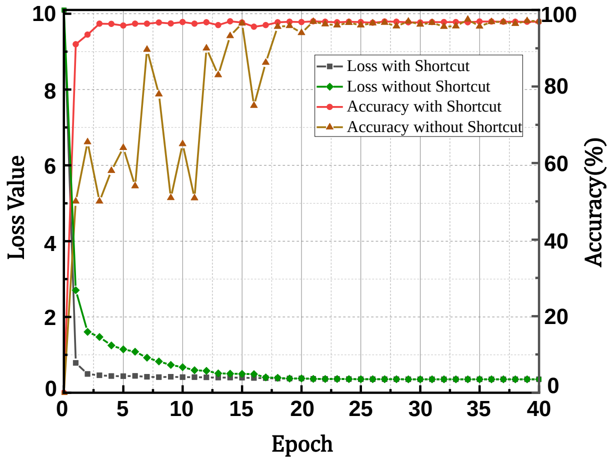
[1]Huazhong University of Science and Technology
[2]State Grid Corporation of China
[3]State Grid Electric Power Research Institute, Wuhan 430074, China
[4]State Grid Corporation of China, Beijing,
[5]Electric Power Research Institute, State Grid Anhui Electric Power Company
[6]School of Electrical Engineering and Automation, Hefei University of Technology
[7]Huazhong University of Science of Technology

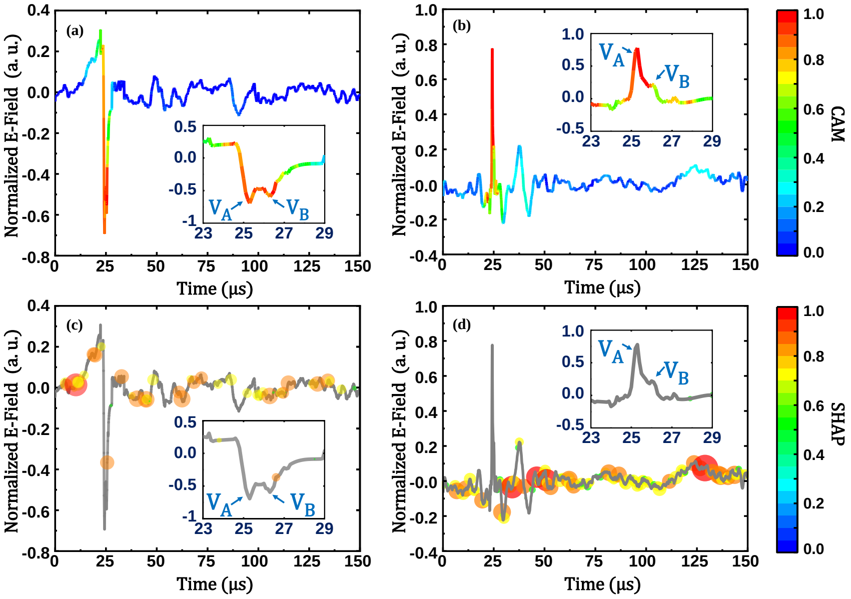
January 24, 2023

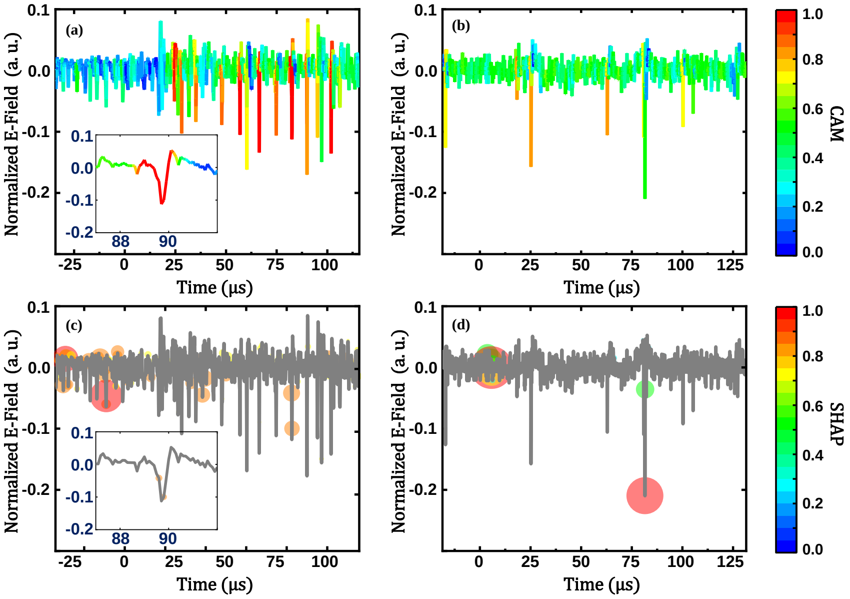
Abstract

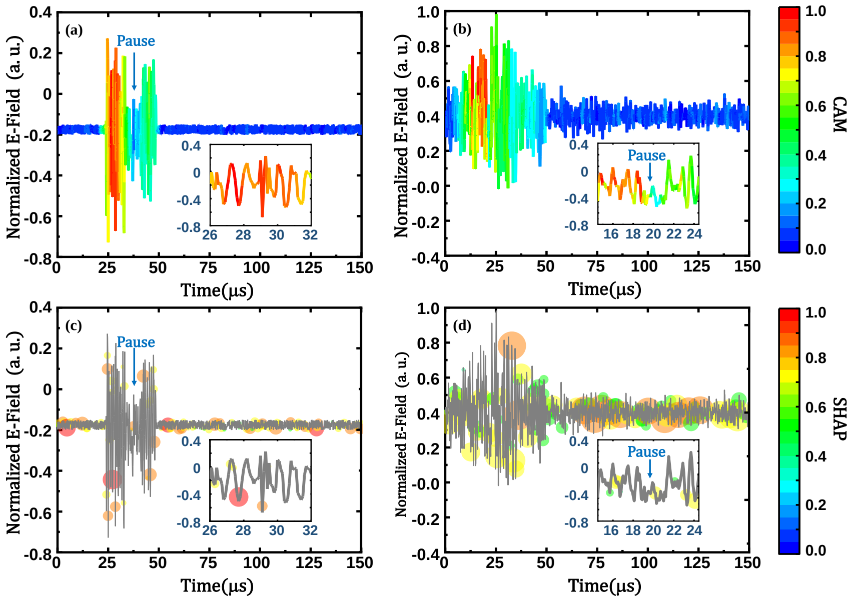
Classification of lightning produced VLF/LF signals plays crucial role in the detection and location of lightning flashes. The machine learning method has potential in the VLF/LF lightning signal classification. Traditional machine learning methods are data-driven and work in a black-box fashion, making the classification accuracy highly dependent on the size and quality of dataset. In this paper, an interpretable convolutional neural network model is proposed for VLF/LF lightning electric field waveform classification. Multi-scale convolutional kernels and shortcut connections are adopted in this model to enhance the ability to capture local waveform features. The CAM method is embedded in our model to open the black-box by visualizing the weight of different waveform features on the classification results. Based on the measured data from five different provinces in China, an accuracy of 98.5% is achieved in a four-type classification task including RS, active stage of IC, PB and NB. The correlation between the weight values of different waveform features and corresponding lightning discharge process are analyzed. It is found that the proposed model can extract decisive features of VLF/LF lightning signals closely related to the physical process of lightning discharges, which is similar to the human expert's behavior. The proposed model is validated by using an open-source dataset from Argentina. It is indicated that the proposed model can resist the impact of unexpected waveform oscillation and achieve a higher accuracy of 98.39% than that of the support vector method. It is demonstrated that our model is less dependent on the training dataset.

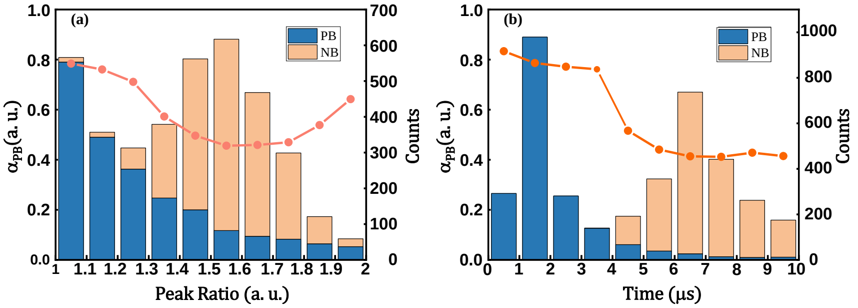

Parameters Update by Backpropagation in Training
step 3
Loss Function
Input Waveform
CNN Model
shortcut
shortcut
step 1
conv. blocks
Conv Layer Structure
i_{th} layer's input
1×1 kernel
10×1 kernel
20×1 kernel
40×1 kernel
3×1 max pooling
1×1 kernel
i_{th} layer's output
Final Layer's Output
$f_i(x)$
$(i = 1, 2, \dots, n)$
Classification
step 2
F^1
F^2
F^3
F^n
α_1
α_2
α_3
α_n
$F^i = \frac{1}{m}\sum_{j=1}^{m} f_i(x_j)$
global average pooling
RS
ASIC
PB
NB
soft max
Pred. C_k
Prob. S_k
results
Physical Information Analysis
step 4
CAM Output
CAM Interpretation
$\alpha_1 \times$
$+\alpha_2 \times$
$+\alpha_i \times$
…
$+\alpha_n \times$
$f_1(x)$
$f_2(x)$
$f_i(x)$
$f_n(x)$
$M_C(x) = \Sigma_k \alpha_k^c f_k(x)$

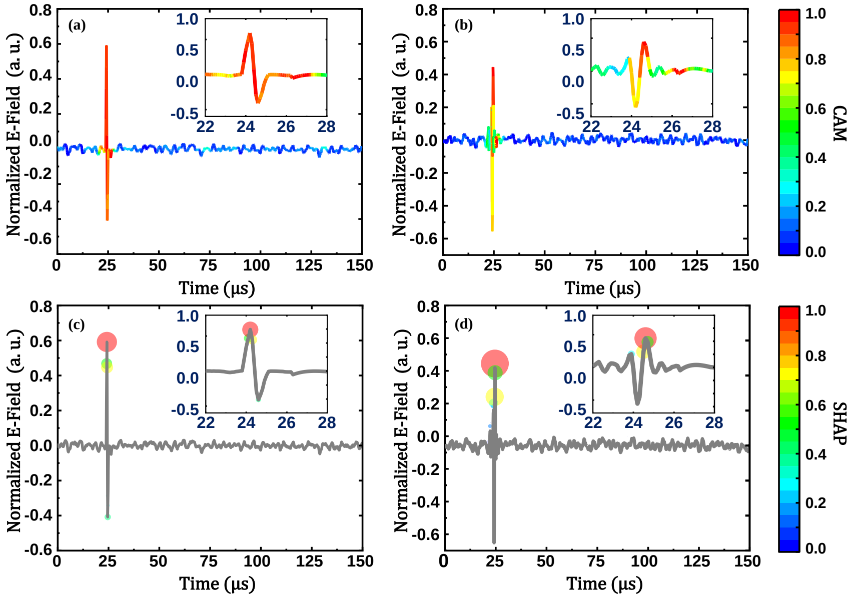

Loss Value
Accuracy(%)
Epoch
0
2
4
6
8
10
0
20
40
60
80
100
0
5
10
15
20
25
30
35
40
Loss with Shortcut
Loss without Shortcut
Accuracy with Shortcut
Accuracy without Shortcut

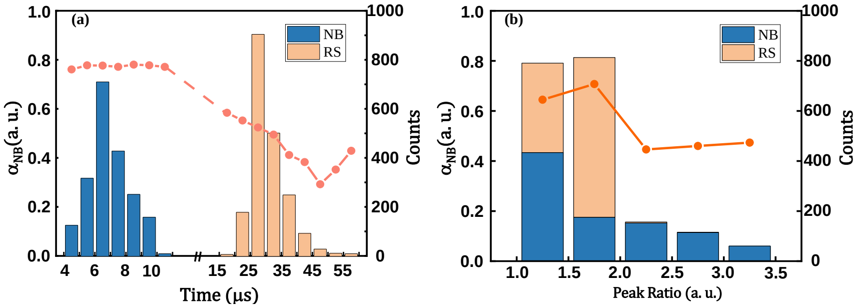

(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (μs)
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
1.0
0.8
0.6
-0.0
0
25
50
75
100
125
150
0.5
0.0
-0.5
-1
23
25
27
29
V_A
V_B
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0

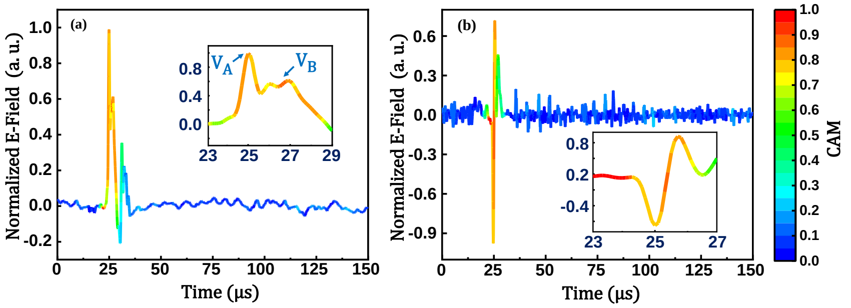

(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (μs)
0.1
0.0
-0.1
-0.2
-25
0
25
50
75
100
125
88
90
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0


(a)
Pause
Normalized E-Field (a. u.)
Time(μs)
0 25 50 75 100 125 150
0.4 0.2 0 -0.2 -0.4 -0.6 -0.8
26 28 30 32
(b)
Pause
Normalized E-Field (a. u.)
Time(μs)
1.0 0.8 0.6 0.4 0.2 -0.0 -0.2 -0.4
16 18 20 22 24
CAM
1.0 0.8 0.6 0.4 0.2 0.0
(c)
Pause
Normalized E-Field (a. u.)
Time(μs)
(d)
Pause
Normalized E-Field (a. u.)
Time(μs)
SHAP
1.0 0.8 0.6 0.4 0.2 0.0


(a)
PB
NB
α_{PB}(a. u.)
Counts
Peak Ratio (a. u.)
1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2
0.0 0.2 0.4 0.6 0.8 1.0
0 100 200 300 400 500 600 700
(b)
PB
NB
α_{PB}(a. u.)
Counts
Time (μs)
0 1 2 3 4 5 6 7 8 9 10
0 200 400 600 800 1000


(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (µs)
CAM
SHAP
0.0
0.2
0.4
0.6
0.8
1.0
0
25
50
75
100
125
150
22
24
26
28


(a)
(b)
NB
RS
α_{NB}(a. u.)
Counts
Time (µs)
Peak Ratio (a. u.)
4
6
8
10
15
25
35
45
55
1.0
1.5
2.0
2.5
3.0
3.5
0
200
400
600
800
1000


(a)
(b)
V_A
V_B
Normalized E-Field (a. u.)
Time (µs)
CAM
0
25
50
75
100
125
150
23
25
27
29
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

[Journal of Geophysical Research Atmosphere]

Supporting Information for

[Towards an Interpretable CNN Model for the Classification of Lightning Produced VLF/LF Signals]

[Lilang Xiao[1], Weijiang Chen[4], Yu Wang[3], Kai Bian[4], Zhong Fu[2], Nianwen Xiang[5], Hengxin He[1], Yang Cheng[2]]

[1 State Key Laboratory of Advanced Electromagnetic Engineering and Technology, HUST, Wuhan, People's Republic of China

2 Electric Power Research Institute, State Grid Anhui Electric Power Company, Hefei, People's Republic of China

3 Wuhan NARI Co., Ltd of State Grid Electric Power Research Institute, People's Republic of China

4 State Grid of China, Beijing, People's Republic of China

5 Hefei University of Technology, School of Electrical Engineering and Automation, People's Republic of China]

Contents of this file

Dataset S1 Waveform data used for the improved CNN model training and test mentioned in this paper.

Data Set S1. The VLF/LF lightning waveform dataset is stored in the .arff format and is available at https://github.com/Massachute/Waveform-data .

Towards an Interpretable CNN Model for the Classification of Lightning Produced VLF/LF Signals

Lilang Xiao[1], Weijiang Chen[4], Yu Wang[3], Kai Bian[4], Zhong Fu[2], Nianwen Xiang[5], Hengxin He[1], Yang Cheng[2]

[1] State Key Laboratory of Advanced Electromagnetic Engineering and Technology, HUST, Wuhan, People's Republic of China

[2] Electric Power Research Institute, State Grid Anhui Electric Power Company, Hefei, People's Republic of China

[3] Wuhan NARI Co., Ltd of State Grid Electric Power Research Institute, People's Republic of China

[4] State Grid of China, Beijing, People's Republic of China

[5] Hefei University of Technology, School of Electrical Engineering and Automation, People's Republic of China

Corresponding author: Hengxin He (hengxin_he@hust.edu.cn)

Key Points:

- The proposed model can extract decisive features of VLF/LF lightning signals which is similar to the human expert's behavior.
- The model achieved an accuracy of 98.5% on a four-type lightning VLF/LF electrical waveforms dataset.
- Testing with data from Argentina validates that the accuracy of the model is less dependent on training data set.

Abstract:

Classification of lightning produced VLF/LF signals plays crucial role in the detection and location of lightning flashes. The machine learning method has potential in the VLF/LF lightning signal classification. Traditional machine learning methods are data-driven and work in a black-box fashion, making the classification accuracy highly dependent on the size and quality of dataset. In this paper, an interpretable convolutional neural network model is proposed for VLF/LF lightning electric field waveform classification. Multi-scale convolutional kernels and shortcut connections are adopted in this model to enhance the ability to capture local waveform features. The CAM method is embedded in our model to open the black-box by visualizing the weight of different waveform features on the classification results. Based on the measured data from five different provinces in China, an accuracy of 98.5% is achieved in a four-type classification task including RS, active stage of IC, PB and NB. The correlation between the weight values of different waveform features and corresponding lightning discharge process are analyzed. It is found that the proposed model can extract decisive features of VLF/LF lightning signals closely related to the physical process of lightning discharges, which is similar to the human expert's behavior. The proposed model is validated by using an open-source dataset from Argentina. It is indicated that the proposed model can resist the impact of unexpected waveform oscillation and achieve a higher accuracy of 98.39% than that of the support vector method. It is demonstrated that our model is less dependent on the training dataset.

Plain Language Summary

Electromagnetic waveforms in very low frequency and low frequency (VLF/LF) band are usually used to detect and locate different lightning activities. Traditional classification methods often misclassify in multi-type lightning discharge waveform classification. The machine learning models show promising potential in the multi-type classification task. However, these models cannot explain which part of the input waveform leads to the classification result, which makes the classification model unreliable. In this paper, we propose an improved and interpretable convolution neural network model, which is adapted to the lightning waveform classification task with changes in model structure. By utilizing the convolution outputs, the model can visualize the contribution of different parts of the waveform to the classification result. The analysis of the visualization results show that the high accuracy and generalization of the proposed model comes from the capture of waveform features corresponding to the key physical process in waveform generation. The dataset for model training comes from five provinces in China, which contains different meteorological conditions. The trained model based on the dataset reached a classification accuracy of 98.5% on test set and 98.39% on another open-source dataset from Argentina, which validated the generalization of the proposed model.

1 Introduction

Remote sensing the electromagnetic radiation generated by lightning discharges is an effective approach to detect and locate lightning activities. It is recognized that the radio emission in the VHF regime is primarily emitted by the streamer and leader involved in lightning discharges, while most of the radiation power is concentrated in the VLF/LF band that is mainly produced by the return stroke (RS) in cloud-to-ground flashes (CGs) and the active stage of intro-cloud flashes (ICs). The detection of VLF/LF radiation was initially introduced to sense the occurrence of CG remotely. Combined with the VLF/LF sensing and the time of arrival (TOA) method, the lightning location system (LLS) was proposed in 1980s which becomes an important technique to support the lightning protection for ground infrastructures nowadays. In order to exclude the impact of ICs, a fundamental task of LLS based on the VLF/LF detection is to recognize the characteristic waveforms produced by return strokes. In recent years, with the development of hardware performance, the emission source location of CGs and ICs can be achieved by using the short-baseline VLF/LF sensing technique and the 3D TOA method. The updated VLF/LF system not only can be utilized as an effective tool for lightning protection engineering applications, but also has the potential in lightning physics research. The lightning leader development process were investigated by using this technique, including the propagation of negative downward leader, the preliminary breakdown (PB), and the narrow bipolar event (NBE) etc. (Bitzer et al., 2013; Y. Wang et al., 2016; Wu et al., 2018). In order to improve the performance of the short-baseline system in lightning detection and lightning physics research, challenges arise in the accurate and automatic identification of waveform characteristics that produced by different lightning discharges.

For most LLS, the multi-parameter method is employed as the criterion to classify the CG and IC, which is derived from extensive field records(Murphy et al., 2021). It adopts specific parameters that can describe the primary profile of VLF/LF waveform, such as the amplitude, the rise and fall time, and the zero-cross time, etc. According to the results of validation studies, the RS detection efficiency of typical lightning location systems (including National Detection Networks (NLDN) and Earth Networks Total Lightning Network (ENTLN) in US, European Cooperation for Lightning Detection network (EUCLID) in Europe ranges from 71% to 92%, while the ICs detection efficiency varies from 73% to 96%(Biagi et al., 2007; Mallick et al., 2015; Schulz et al., 2016). Despite the difference in hardware performance, the deviation in detection efficiency of different systems is mainly attributed to the classification accuracy of CGs and ICs. On the one hand, since the multi-parameter method is difficult to extract characteristic parameters from VLF/LF signals with low-amplitude, the small signals were often abandoned, resulting in the decrease of detection efficiency(Kohlmann et al., 2017; Nag et al., 2014). On the other hand, the characteristic parameter involved in the multi-parameter method may vary in regions with different meteorological conditions(Cooray, 2009; Said et al., 2010; Shao & Jacobson, 2009; Wooi et al., 2015). For instance, the rise time and zero cross time of RS in Vitemölla, Sweden is of 5-25μs and approximately 40μs respectively, while the rise time decreases to about 2.5-9μs and the zero cross time increases to the range of 40-160μs in Sri Lanka(Cooray & Lundquist, 1982, 1985). Accurately determining the thresholds of characteristic parameters requires the support of long-term data. Recently, the machine learning methods such as the support vector machines (SVM) and the convolutional neural networks (CNN) are introduced to improve the classification efficiency of lightning VLF/LF signals. The SVM method is utilized to classify the VLF/LF lightning waveforms of CGs and ICs. A classification accuracy of 97% is achieved, which shows an excellent adaptability and automation(Zhu et al., 2021). The CNN models with different structures are proposed to perform the classification of VLF/LF signals generated by multiple lightning processes, including RS, PB, and NBE, etc. (Peng et al., 2019; J. Wang et al., 2020). It indicates that CNN has the potential to realize signal classification produced by various complex lightning discharge processes.

Although extensive efforts have been paid to improve the classification accuracy of lightning VLF/LF waveforms, towards to the development of high-performance short baseline VLF/LF lighting detection system, the following limitations still exist:

- Using the multi-parameter method, the classification accuracy of RS and IC in the LLS system has reached more than 90%. The classification accuracy may be further improved by optimizing thresholds of the multi-parameter method based on long term operation experience. However, since the VLF/LF waveforms produced by lightning leader discharges has more pulses and other high frequency components, it is difficult to determine thresholds involved in the multi-parameter method which can effectively discriminate different lightning events correlated to lightning leader propagation. Recently, it was found that the VLF/LF signals generated by NBE are wrongly identified as RS by the multi-parameter method(Leal et al., 2019; Lyu et al., 2015).
- The machine learning methods show promising performance in multi-object classification tasks, the challenges of applying machine learning methods in lightning VLF/LF waveform classification come from two aspects. Firstly, note that the data-driven nature of the machine learning methods means that the

performance is highly dependent on the balance and quality of the original dataset. An CNN model derived from imbalance data set is not reliable, because the model will tends to classify the objective waveform into the category which has the most samples in training dataset (Kaur et al., 2019). Secondly, since the characteristics of lightning VLF/LF signals can change in different regions, the accuracy of machine learning methods largely depends on whether the training dataset covers all possible variations of the objective waveform characteristics. Meanwhile, we need to note that most of the classification process by using machine learning methods acts like black box models, which makes it is difficult to ensure the classification accuracy of different lightning events. As discussed by Zhu et al., misclassification of RS signal can still occur by using SVM, although the characteristics of the misclassified waveform can be easily recognized manually. Since it is difficult to obtain the lightning waveforms in all regions of the world to expand the database, it is necessary to develop interpretable machine learning models to open the black box, which can reveal the classification process (Lipton, 2018) and assess whether the model is able to capture the essential characteristics of different types of lightning VLF/LF signals.

In this paper, a new interpretable CNN model which utilizes the class activation map (CAM) to represent the contribution of different waveform parts during the classification process is proposed. A four-class dataset including RS, PB, NB and IC is established for model training. The dataset is based on 17,441 waveforms recorded from five provinces in China with the latitude ranging from 29.1° to 33.5° and the longitude from 91.1° to 120.2°. The classification accuracy of the trained CNN is compared with that of the SVM model. The classification process of the four types of lightning waveforms is visualized by the CAM, which throws light on the relationship between the high-weighted waveform features and the physical process of leader discharge in lightning. The classification results are analyzed for the range of variation of the characteristic parameters of different waveforms in turn. The generalization of the proposed CNN model is test on another open dataset in Argentina used by Zhu et al. This paper is organized as follows: Section 2 introduces the data sources and the improved CNN network structure used in this paper. Section 3 shows the classification performance of the trained model and discusses the interpretability of the classification results. Section 4 discusses the universality of the CNN model, and Section 5 makes the conclusion.

2 Data and Methodology

2.1 Dataset

The dataset used in this paper comes from 17,441 lightning radiation waveform data recorded during 2019-2020. The measurement device is the VLF/LF electric field change meter (EFCM). The EFCM consists of an antenna and a digital data acquisition unit. The frequency band of the EFCM is 10Hz to 500kHz, with a sampling rate of 5Ms/s and a GPS synchronization error of less than 50ns (Y. Wang et al., 2020). The EFCMs were installed in five different provinces of China, including Hubei, Jiangsu, Zhejiang, Anhui and Tibet, as shown in Table 1. When the dataset covers waveforms in a variety of meteorological and terrain conditions, the model can be more generalized and the classification accuracy may be improved. The installation sites of these EFCMs have altitude between 190m to 4000m above sea level, within the longitude from 91.1° to 120.2° and the latitude from 29.1° to 33.5°. In order to improve the quality of recorded waveforms and exclude the impact of measurement noises, a combination of empirical mode decomposition (EMD) method and wavelet denoise method were used to pre-process the lightning radiation signal.

Table 1 Location of the deployed VLF/LF lightning waveform measurement meters

Location	Longitude	Latitude
Wuhan, Hubei	114.409	30.514
Sihong, Jiangsu	118.219	33.482
Wuxi, Jiangsu	120.256	31.618
Lishui, Zhejiang	119.656	27.976
Taizhou, Zhejiang	121.38	29.125
Hefei, Anhui	117.202	31.761
Anqing, Anhui	116.123	30.231
Lasa, Tibet	91.14	29.666
Linzhi, Tibet	94.373	29.636
Changdu, Tibet	97.179	31.146

Compared to the multi-parameter classification method, the machine learning algorithms can utilize the entire data information of time-resolved waveforms instead of several characteristic parameters. The SVM method was employed to classify the full VLF/LF waveform of lightning signals (Zhu et al. 2021). Considering the computational resources required for the deployment, the waveform is down sampled and divided into equal lengths, where each waveform slice is 100 μs in duration and contains 101 sample points. With the development of microprocessor technology in recent years, the main frequency of LS1043A ARM board we use has reached 1.6GHz, which significantly improves the ability to process waveform data per unit of time. In this paper, each waveform slice contains 2500 points corresponding to a time duration of 500 μs, which is beneficial to preserve the essential features of the original waveform. In the following discussion, each waveform slice is called as a sample. The dataset is constructed based on the prior knowledge of RS, PB, NB and active stage of IC, which can be found in reported literatures. We manually selected samples with the highest signal-to-noise ratio (SNR) and divided them into these four categories with a total of 8000 samples. To ensure a balanced dataset, each of the four categories contains 2000 samples. Note that our dataset is not classified by the polarity of lightning event, because the polarity can be easily intensified according to the polarity of the first pulse. It should be emphasized that this simplification will not affect the classification results discussed in the following parts. A 5-fold cross validation approach was adopted for the dataset, with the training set containing 6000 samples and the rest 2000 samples participate in the test.

2.2 Method

The CNN model performs feature extraction by convolving a convolution kernel with the input data. The convolution kernel is a weight matrix representing the features learned by the CNN model. The convolution kernel is usually initialized with random values, and the CNN model compares the output results with the true labels and updates the convolution kernel by backpropagation during the iterations. Thus, the convolution kernel can better match the core features of the data and improve the model's performance. When making VLF/LF lightning waveform classification, traditional CNN networks (plain CNN) have the following limitations:

- The same size of the convolutional kernel in each convolutional layer in plain CNN makes it difficult to handle the possible multi-scale features in waveforms.
- We use a high computing capability development board to process the data, the sample time in the dataset is longer and the sample includes more information, requiring more convolution layers to fully extract these features. However, the backpropagation process calculates derivatives in chains to update network parameters. The gradient information may vanish gradually when simply increasing the number of convolution layers. It can also cause model degradation and make the model difficult to converge (He et al., 2016).
- The CNN model flattens the features extracted from the convolutional layers into a one-dimensional vector, and output the classification results by means of full connection. It is difficult to obtain which part of the waveform determines the classification results of the model, which makes it impossible to judge whether the

classification process of the model is reliable.

In order to fix those issues, we proposed an interpretable CNN model with improved performance in feature extraction and convergence speed. The model includes a CNN classifier and a visualization module as shown in Figure 1.

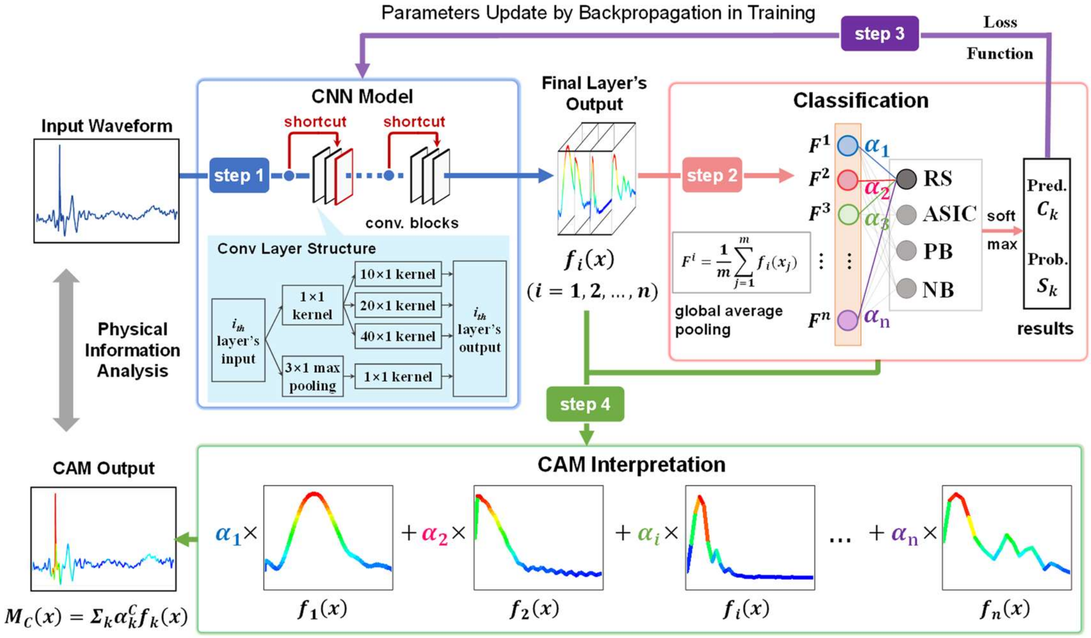


Figure 1 Structure of the proposed interpretable CNN model

a) ***The CNN classifier:*** The proposed CNN classifier takes waveforms as input and gives out classification results with probabilities. Step 1-3 describes how the CNN classifier works and self-upgrades iteratively in training.

In step1 the waveform is fed into the CNN model and the high-dimensional feature maps are obtained. Compared with the plain CNN, the proposed CNN model adopts shortcut connections and parallel convolution kernels. The CNN model contains two convolution blocks, which is formed by stacking three convolutional layers. In each block, part of the input data is directly transferred to the second layer of the block through a shortcut connection. The shortcut connection aims to solve the problem of model degrading in multi-layer networks and accelerates the convergence in training. In each convolution layer, the convolutional kernels with the size of 40, 20, 10 and 1 are introduced in a parallel structure. The kernels with the size of 40, 20 and 10 give the model a more various feature matching range after multiple layers, enabling the extraction of long-scale waveform features. The kernels with the size of 1 ensure that the model can also capture detailed waveform features. Each layer can be expressed as:

$$f^l = b^l + \sum_{i=1}^{N_{l-1}} conv1D(w^l, f_i^{l-1}) \qquad (1)$$

Where x^l is defined as the input of layer l, b^l is defined as the bias layer l , f_i^{l-1}

is the i_{th} output part of layer *l-1*, w^l is the multi-size convolution kernels at layer *l*, N_{l-1} is the number of output in layer *l-1*.

In step2, the feature vector is obtained through the global average pooling based on the feature maps produced in step 1. Compared with the plain CNN methods, which flatten the high-dimensional feature maps as feature vectors, the proposed model uses the global average pooling to form the feature vectors and greatly reduces the computations of the model.

The classification probability of the waveform is computed by the fully connected layer and the SoftMax function. The probability S_c that a waveform belongs to a category *c* can be obtained from equation (3):

$$S_c = \Sigma_i \alpha_i^c F^i \qquad (2)$$

Where α_i^c represents the contribution of feature map $f_i(x)$ to model's classification result of category *c*.

During the model training, the model's classification will be compared with the true label of the waveform by the loss function as shown in step 3. The result is referred as the loss value in training. The model uses the back propagation algorithm to make the loss information flow backward to update model parameters, which can be expressed as:

$$\frac{\partial u^n}{\partial u^j} = \sum_{i:j\in Pa(u^i)} \frac{\partial u^n}{\partial u^i}\frac{\partial u^i}{\partial u^j} \qquad (3)$$

Equation 5 describes how to calculate the gradient of an output node u^n (such as the loss value) over several input nodes from u^1 to u^j to achieve gradient descent update of the parameters. Where u^i refers to the intermediate nodes in all possible paths (Pa) from u^n to u^j. The gradient is essential for the gradient descent optimizing method during parameters update.

b) ***Model Interpretation:*** In order to open the black-box of the CNN model, we introduce the CAM method in the proposed CNN model as shown in step 4. The CAM method multiply the weight vector produced in step 3with the high-dimension feature maps produced in step 1 to obtain a class activation map (CAM) which can mark the important waveform features in the classification. The CAM values for the class can be defined as：

$$M_C(x) = \Sigma_i \alpha_i^C f_i(x) \qquad (4)$$

We denote $M_C(x)$ as the CAM value of the waveform under category *c*. By using heat maps the CAM value provides a direct indication of the importance of each datapoint *x* to the classification result of category *c*.

Due to the model structure difference, CAM method is not appliable on traditional machine learning methods like the SVM. We use the SHAP method to visualize the important waveform features in the classification of traditional machine learning methods for comparison. SHAP method derives from cooperative game theory, which provides global and local interpretability of the features(Lundberg & Lee, 2017). The SHAP value is based on the marginal contribution of the features amongst all the feature arrangements. In waveform classification, we regard each waveform datapoint as a feature, the SHAP value can be expressed as:

$$\varphi_j(val) = \Sigma_{S \subseteq x_1,\cdots,x_M/x_j} \frac{|S|!\,(M-|S|-1)!}{M!}\left(val\left(S \cup x_j\right) - val(S)\right) \qquad (5)$$

Where the x_j is the j_{th} feature and refers to the j_{th} point of input waveform. φ_j is the contribution value of x_j to the classification result. S is the subset of features and M is the total number of features. The value function $val(*)$ refers to the model's classification results in machine learning. In the following section, SHAP is referred to the φ_j and used to describe the contribution of different waveform features to the classification results in traditional machine learning methods.

2.3 Model Training

In this paper, the proposed CNN model is deployed on a Tesla A100 graphics card and the training framework is Pytorch 1.9.1. The hyperparameters required for training are shown in Table 2.

Table 2 Hyperparameters used for training

Hyperparameter	Value
Batch Size	48
Epoch	40
Loss Function	CrossEntrophy
Optimizer	Adam
Learning Rate	0.01
Momentum	0.5

Batch Size means that the model is fed with 48 data samples at a time in model training, and the parameter Epoch specifies that the model performs a total of 40 forward calculations and back propagation processes. The loss function used in the model is the CrossEntrophy function and the Adam is used as the optimization algorithm. The Learning Rate and Momentum together control the convergence rate and the efficiency of the model, which are set to 0.01 and 0.5 respectively after pre-test. The model was trained under the above conditions, and the training loss and classification accuracy are shown in Figure 2. Due to the introduction of the shortcut connection to solve the vanishing gradient problem, the model converges faster and reaches convergence after only five epochs, with an overall classification accuracy of about 98.5%.

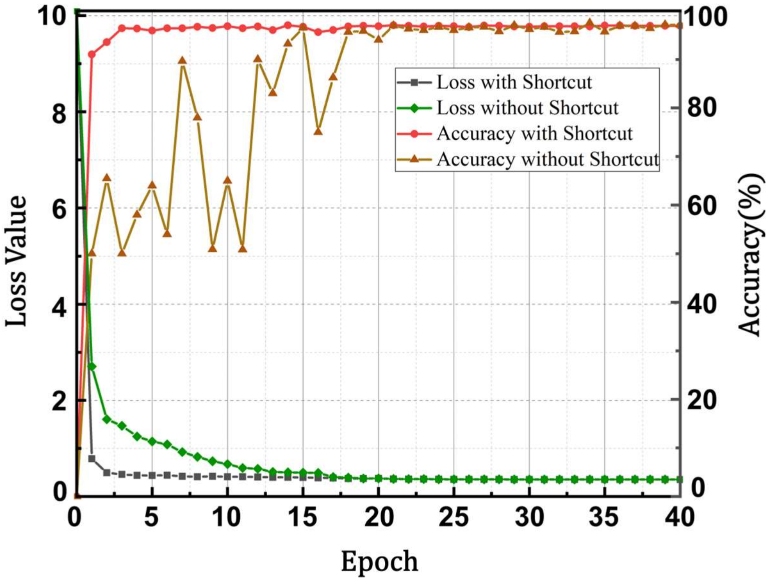


Figure2 Loss and accuracy changes of plain CNN model without shortcut and improved CNN model with shortcut in training

3 Results

3.1 Comparison for classification results

After training, we compared the classification results of the proposed CNN method with other machine learning methods such as SVM and RF under the same dataset. The feature vector used for training SVM is obtained by data down sampling method (Zhu et al., 2021) and the amplitude-frequency features were extracted as feature vectors when training the RF model (Nassralla et al., 2017) . The performance of these methods is shown in table 3:

Table 3 Comparison of the results of different models

Method		CNN				SVM	RF
Metrics		Accuracy	Precision	Recall	F1	Accuracy	Accuracy
Class	RS	96.8%	1.00	1.00	1.00	90%	88.3%
	PB	100%	0.94	0.97	0.96	91%	83%
	IC	100%	1.00	1.00	1.00	86.20%	84.5%
	NB	97.8%	0.97	0.94	0.96	92%	92.7%

Table 3 shows the comparison of classification accuracy in the four kinds of waveforms. For waveforms with short duration such as RS and NB, both CNN method and traditional machine learning methods achieve good classification results in classification accuracy, while the CNN model has an improvement of about 6%. However, for waveforms like PB and IC which last longer and is more difficult to classify, the CNN method shows a significant improvement of up to 17% in accuracy compared to traditional machine learning methods. The performance of SVM and RF is not as expected as that in the original literature. This may be due to the fact that the dataset we used contains longer slices of waveform which can preserve more waveform features. Therefore, the hyperparameters used in the original literature may be no longer suitable and need to be adjusted. In contrast, the CNN model supports raw waveforms as inputs, without the need of manual adjustment of hyperparameters after changing datasets, which can bring stronger robustness. In addition, the CNN model performs well on metrics like precision, recall and F1, proving that the model does not have an imbalance problem.

3.2 Model interpretability analysis based on CAM visualization

The proposed CNN model achieves higher performance in all kinds of waveforms, which may be related to a better understanding of the physical process. To investigate the feature the model has learned, we use CAM visualization method mentioned in section 2.2 to estimate which part of the waveform lead to its classification result. For comparison, the SHAP method is used to mark the waveform which owns higher contribution to affect the classification results in SVM(Ribeiro et al., 2016). In this section we will discuss the classification process of the four typical VLF/LF waveforms in both cloud to ground flashes (CG) and intracloud flashes (IC), including return stokes (RS), active stage of IC (ASIC), preliminary breakdown (PB) and narrow bipolar pulses (NB). The accurate classification of RS helps to distinguish between CG and IC events and can improve the location efficiency of LLS. The classification of PB, NB and ASIC is important for further research of the initiation mechanisms of the lightning discharges.

3.2.1 Return Stroke Produced by CG

The LLS allows for real-time detection of return strokes (RS) in cloud to ground flashes (CG) due to the strong and widely spread VLF/LF waves of RS, which is generated by the propagation of high amplitude currents in the lightning channel. The waveshape and amplitude of RS are closely related to factors like current strength, propagation speed and propagation path etc. Compared to other stages of the lightning discharge, the velocity of RS is high and the current of RS is strong. Therefore, the LLS often use the amplitude or pulse width of the electrical field waveform to identify RS. The traditional multi-parameter classification method concludes that the pulse width of the RS waveform is typically between 10 and 200 us. However, recent reports indicate that using pulse width as a criterion can easily misclassify several kinds of intracloud lightning pulses as RS (Biagi et al., 2007; Leal et al., 2019; Nag et al., 2014). In recent years, scholars have already tried using machine learning methods like SVM to perform RS/IC classification (Zhu et al., 2021). This section applied the CNN model trained in Section 3.1 for RS classification. By analyzing the classification process of our CNN model under several typical RS cases, we explored the key physical features the model learned and compared it with the SVM method.

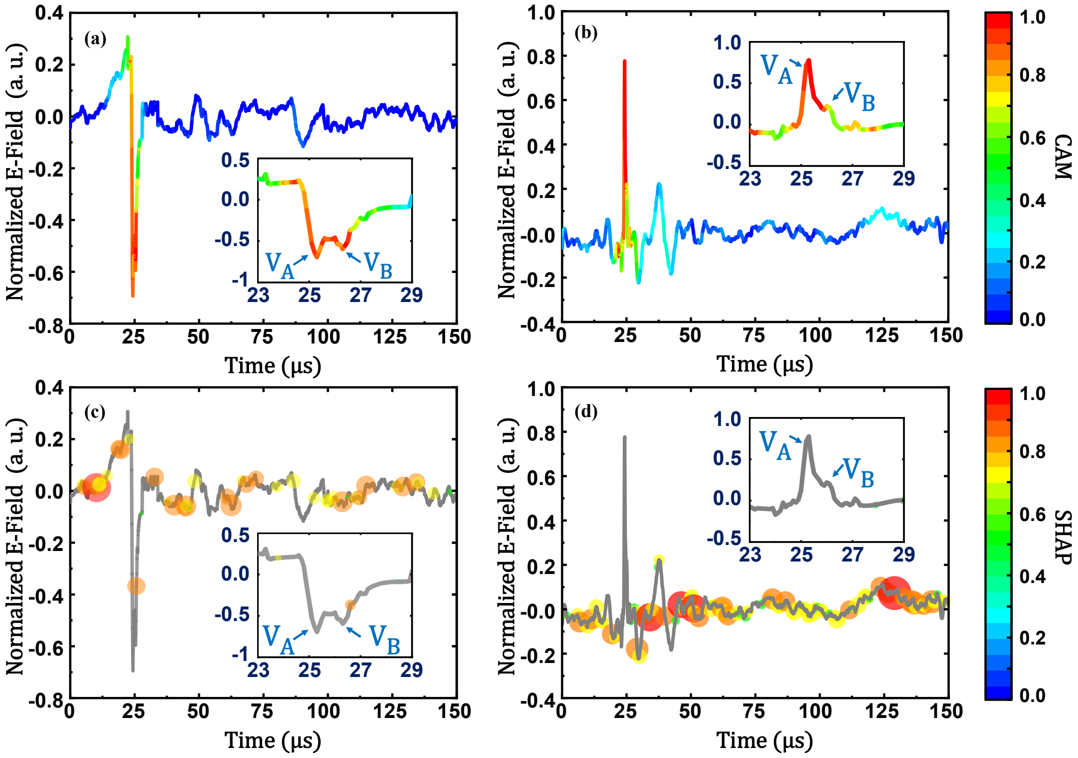


Figure 2 Visualized classification results of negative RS waveform #190912135803-001RS and positive RS waveform #190912135803-003RS waveform. **a)** Visualized CNN classification result based on CAM for the negative RS case with a detailed demonstration for the main pulse part **b)** Visualized CNN classification result based on CAM for the positive RS case with a detailed demonstration for the main pulse part **c)** Visualized SVM classification result based on SHAP for the same case and detailed demonstration as (a) **d)** Visualized SVM classification result based on SHAP for the same case and detailed demonstration as (b)

Figure 3(a) shows the classification result of the CNN model for a negative RS, which was recorded at 13:58:03 September 2019 at Anqing, Anhui, China. We define data points with CAM weight values above 0.5 as hotspots during classification. The pulse width of this case is about 11 μs, which is at the lower threshold according to the multi-parameter method, leading to great possibility for misclassification. However, the CNN model gives out a hotspot region between 25μs and 27μs, which means the CNN model accomplished the classification mainly by the main peak part with a duration of only 2μs. It can be seen from this part that the main pulse contains a sequential double-peak characteristic with a primary peak VA and a subpeak VB. Based on observation results, Le Vine et al. conclude that the subpeak structure of the RS waveform is related to the geometry change of the lightning channel (Le Vine, 1980). Cooray et al. propose that the abrupt changes in current amplitude or channel development velocity will result in the subpeak structure in the VLF/LF waveform (Cooray & Lundquist, 1985). Figure 3(a) demonstrates that the CNN model successfully captured the double-peak characteristic of the RS waveform, which represents a key part of the physical process of RS.

Figure 3(c) gives the SVM classification result for the same negative RS waveform. The orange and red circles represent the high weight points given by the SVM with a SHAP value greater than 0.5. The high weight points distribution shows that the SVM model is able to depict the profile of the waveform. It can be inferred from this that the SVM method may classify the RS waveform by marking the high-amplitude part of the RS waveform. Although the SVM model was also able to correctly classify this RS event, it failed to capture other physical information of the RS waveform.

The classification results for another positive RS are given in Figure 3(b) and Figure 3(d). The hotspot region in Figure 3(b) shows that the CNN model also captured the two-peak feature of both PA and PB, which means a better understanding about the correlation between the RS waveform and physical process such as the change in the channel development velocity and the channel geometry. However, the SVM method fails to classify this positive RS waveform. As can be seen from Figure 3(d), the high weight points given out by the SVM model also tends to depict the entire waveshape. But due to the bipolar pulse following the main pulse of this waveform, the SVM method failed to mark the main pulse, which resulted in a 42% probability for RS while an 82% probability for NB. The CNN model proposed in this paper not only marks the true main pulse part of the RS waveform but also captured the double-peak feature in the main pulse. Compared to multi-parameter method, the proposed CNN model can overcome the problem that an applicable general criterion for parameters like pulse duration is difficult to be determined. As for traditional machine learning method like SVM, the anti-disturbance capability of the proposed CNN model also gets improved, which means stronger robustness.

3.2.2 Active Stage of ICs

Intra-cloud lightning discharge (IC) occurs in a single storm cloud or between different storm clouds. The intra-cloud lightning discharge can be divided into wo stages, including the active stage and the final stage (Bils et al., 1988). The VLF/LF radiation signal generated during the active stage of intra-cloud lightning (ASIC) presents a sequence of pulse activities. A variety of transient processes appears in the following final stage, including the narrow bipolar events, the stepped leader, the J process and the K process, etc. (Rakov & Uman, 2003). The VLF/LF radiation signal of the final stage of ICs is usually not used to identify the intra-cloud lightning events, because it owes an overlapping amplitude range with the RS. Conversely, during the ASIC, repetitive VLF/LF electric field pulses can be detected. These pulses are characterized by low amplitude and unipolarity and are related to the stepped growth of the negative leader, which are applicable to distinguish the IC and CG(Brunner, 2016). However, the statistics of characteristics of the pulses during the ASICs are rarely reported. In this section, attempts were made to demonstrate that an interpretable CNN model can be used as an effective approach for the classification of ICs events.

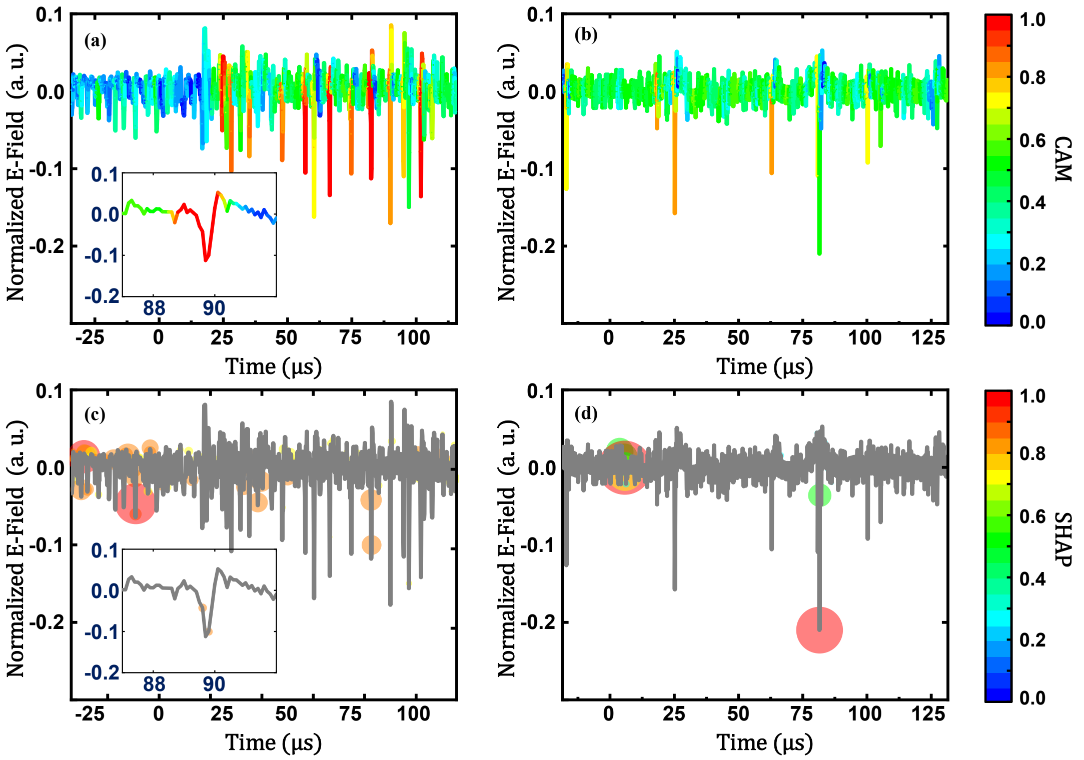


Figure 3 Visualized classification results of negative IC waveform #190818114117-001IC and #190817184223-001IC. **a)** Visualized CNN classification result based on CAM for the first case with a detailed demonstration for the one single pulse **b)** Visualized CNN classification result based on CAM for the second case with a detailed demonstration for one single pulse **c)** Visualized SVM classification result based on SHAP for the first case **d)** Visualized SVM classification result based on SHAP for the second case

Figure 4(a) shows the visualized CNN classification result for a pulse train during the active period of the IC lightning. This waveform is captured at 11:41:17 18th August 2019 at Lishui, Zhejiang, China. The electrical waveform at this stage consists of a sequence of pulses. The tail of each single pulse is followed by a small, slowly changing polarity-reversed process (Krider et al., 1975). The pulses repeat slowly at this stage, with an average pulse interval of 10.7μs in this case, which is consistent with the existing observations (Gomes & Cooray, 2001; Krider et al., 1975). In Figure 4(a), the hotspot region given out by the CNN model mainly contains two parts. Firstly, the model focuses on the pulse peak and its subsequent polarity-reversed part. The mean CAM value of this region is greater than 0.8. It shows that the CNN model provides a good understanding of the causal relationship between the main peak and the subsequent polarity-reversed waveform and treat them as a whole part. Secondly, it should be noted that the waveform between the two pulses is mainly background waveform, but the model still gives out CAM values of 0.5 to 0.6, significantly higher than the CAM values of the background waveform outside the active stage. It can be inferred that the model not only captures the characteristics of single pulses of ICs, but also find the pattern of continuously pulse repeat during the ASIC. In Figure 4(c), the SVM method is used to classify the same case. The circles in Figure 4(c) represent the high weight points given out by the SVM with a SHAP value greater than 0.5. The high weighted points mainly locate near the peak of the pulse. This suggests that although the SVM method can also capture the peak of the pulse, it fails to capture the causal relationship of main peak and the succeeding part.

Figure 4(b) and Figure 4(d) compare the classification results of CNN and SVM method for another IC waveform. For the CNN model, the hotspot region is similar to that in Figure 4(a) which also concentrate on the single pulses and the pulse intervals. However, the SVM method misclassifies the waveform as RS. Figure 4(d) shows that the high weighted points of the SVM model mainly locates around 0μs and 75μs. The waveform around 0μs is mainly the background electric field and the waveform around 75μs has the highest pulse peak. The results show that the SVM model mainly focus on the peak of the electric field pulses. Besides, since the SVM model mainly focuses on the pulse peaks, it makes the high weight points in this case locate around the largest pulse, leading to the misclassification as RS. By the comparison, it can be concluded that the CNN model is able to learn the detailed features like the temporal relationship between the first peak and subsequent polarity-reversed part of pulses, and is also able to effectively identify the macro features like the repetition of pulses in active stage of ICs.

3.2.3 Preliminary Breakdown Pulses

The preliminary breakdown (PB) is the initiation and development of the leaders in cloud, which is considered to be the initial stage of the lightning. The VLF/LF electric field waveform generated by PB is composed of consecutive bipolar pulses with a total duration of microseconds. It is concluded from theoretical simulations that the waveform of the PB process has a similar physical mechanism to that of the NB, which is probably related to the consecutive stepped elongation of the negative leader channel within the thundercloud(Da Silva & Pasko, 2015). The multi-parameter method usually utilizes the SNR to make classification. It is considered to be a PB process while at least three consecutive bipolar pulses are found with peaks twice the average noise level or more (Nag & Rakov, 2008). However, according to the discussion in section 3.2.2, the waveforms during the ASICs are also characterized by repetitive bipolar pulses, which makes it difficult to achieve an accurate distinction between the PBs and ASICs using multi-parameter method. In this section, the visualized result of the CNN model for PB classification is analyzed to illustrate the CNN model's ability to capture the physical features of PB waveforms, which improves the model's accuracy and robustness.

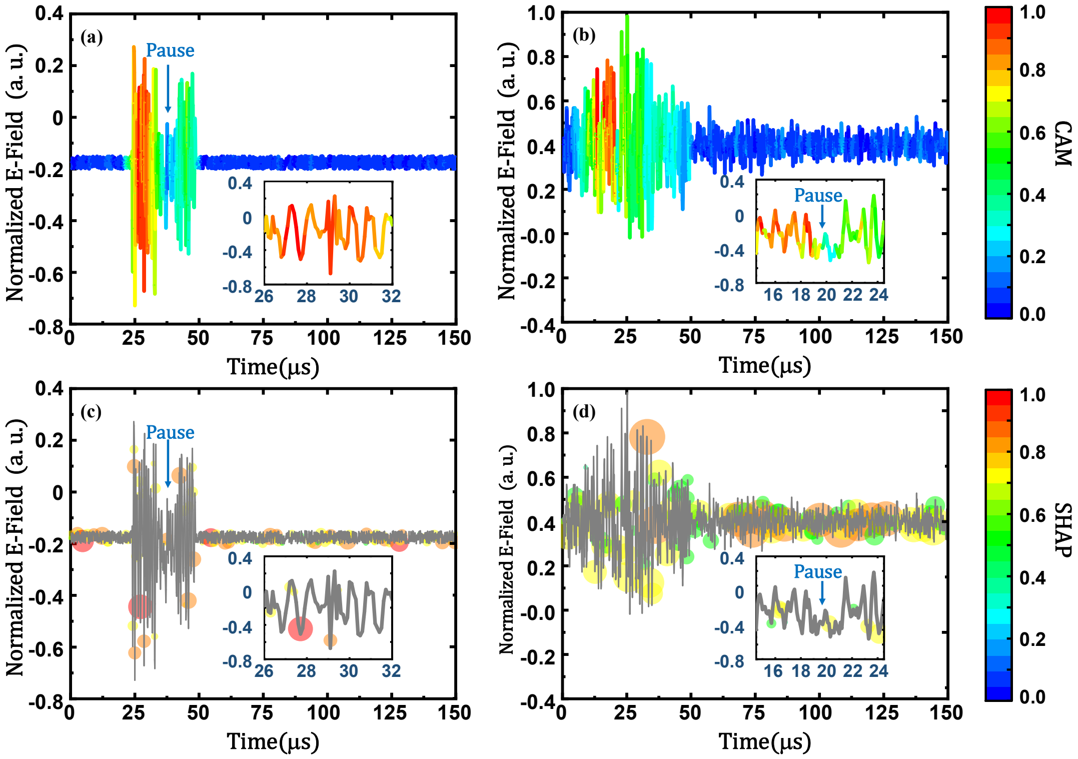


Figure 4 Visualized classification results of PB waveform #190727133157-001PB and #1907271314423-001PB. **a)** Visualized CNN classification result based on CAM for the first PB case with a detailed demonstration for the main pulse and the pause part **b)** Visualized CNN classification result based on CAM for the second PB case with a detailed demonstration for the pause part **c)** Visualized SVM classification result based on SHAP for the first case **d)** Visualized SVM classification result based on SHAP for the second case

Figure 5(a) shows the visualized CNN classification result for an PB waveform, which is captured at 13:31:57 27th July 2019 at Wuxi, Jiangsu, China. The hotspot region given out by the CNN model covers the entire period between 23 and 54μs in which the pulses exist. However, it is obvious that two partitions with different CAM values exist during this period. The first partition is between 26 and 32μs where the CAM values are all above 0.8. It can be seen from Figure 5(a) that the waveform within this partition is notably characterized by significant continuous bipolar oscillations. The second partition is between 38 and 54μs where the CAM values range between 0.4 and 0.6 and also has significant continuous bipolar oscillations. It is interesting that in the time from 32 to 38μs between the two partitions, there is a pause interval where the CAM values are less than 0.3. Figure 5(a) shows that the amplitude is low and bipolar oscillation is not obvious during the interval. The above analysis shows that the model in this paper is able to adaptively mark intervals that match the PB characteristics based on the bipolar oscillation frequency and amplitude characteristics of the waveform. Figure 5(c) shows the visualized SVM classification result for the same case. It can be seen from Figure 5(c) that the high weight points distribute on both the background waveform and the pulse part, which indicates that the SVM completes the PB identification by depicting the overall profile of the waveform without an understanding of PB's core physical features.

To further compare the proposed CNN model with the SVM model, another case of PB is shown in Figure 5(b) and Figure 5(d). In Figure 5(b) the CNN model gives a similar distribution of hotspot regions as in Figure 5(a). The CNN model accurately marks the pulse part which is also divided into two partitions by a pause interval, and the duration of the pause interval in Figure 5(b) is shorter by about 1μs compared to figure 5(a). This represents a better ability of the CNN model to adaptively classify continuous pulses that conform to bipolarity, with better robustness. The above phenomenon means that the CNN model is able to adaptively find PB-like waveforms, and even a very short non-PB interval will cause the CAM value to drop sharply. In comparison, Figure 5(d) shows that the SVM method misclassifies this case, with the highest weight points distributing at the start of the waveform around 30μs. Besides, the other high weight points mainly exist in the negative part of the waveform, leading to the misclassification. It can be inferred that the CNN model has higher classification accuracy compared to traditional machine learning methods for the ability to recognize temporal features like the continuous bipolar pulses.

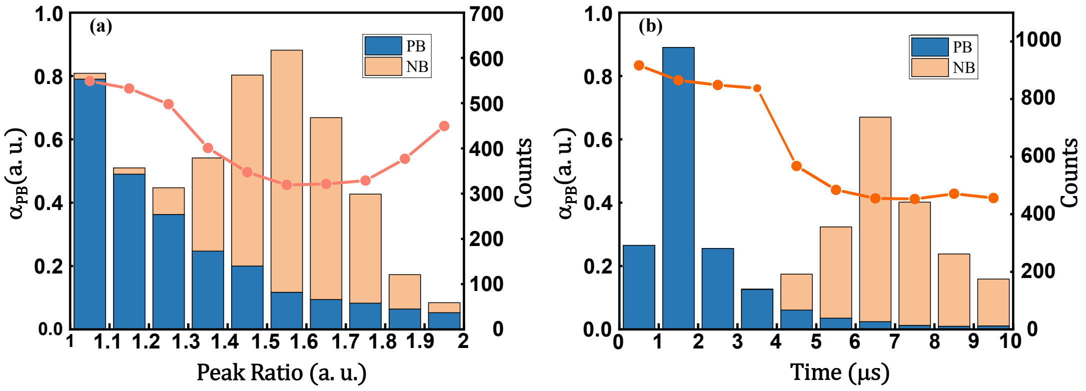


Figure 6 (a) relationship of CAM values α_{PB} and peak ratio P (b) relationship of CAM values α_{PB} and pulse duration T

To demonstrate our model's ability to help to find adequate threshold for multi-parameter classification. The average CAM values α_{PB}, peak ratio P and pulse duration T are estimated and compared with the results of similar waveforms like NB. We define a single pulse as a segment of the waveform between two zero crossing points, which must contain a polarity change. The α_{PB} refers to the average CAM value given by the model when the waveform is considered to be a PB. The peak ratio P is the absolute value of the first and second peak amplitude ratio, and the pulse duration T is the time interval between two crossing points. The α_{PB} -P relationship is given in Figure 6(a). The result shows that the bipolar peak ratio of the PBs waveform ranges between 1 and 2, which is in agreement with the range of peak ratios of the NB. The difference is that 63% of the PBs have a P of less than 1.3, while 81% of the NBEs have a P between 1.3 and 1.7. It is notable that when the P is less than 1.3, the α_{PB} is greater than 0.5 and decreases as the P increases. This suggests that the more P is close to 1, the more likely the pulse considered to be PB in our model, which is consistent with the simulation results of Silva et al. (Da Silva & Pasko, 2015). However, when the P is greater than 1.8, the α_{PB} increases as the P increases. This suggests that the peak ratio cannot be used as an effective way to distinguish PB from NB. This may be because that the P changes as the conductivity of the leader channel changes in which the PBs radiation source locate, thus deviating from the theoretical result of Silva(Kašpar et al., 2017). Figure 6(b) demonstrates the relation of α_{PB} -T, where T is within 10 μs for both PB and NB. 91% of PB had T of less than 4.0 μs, while all of NB have T between 4.0 and 10.0 μs. As can be seen from the trend of α_{PB}, the proposed model suggests that the shorter the pulse duration is the more likely the pulse is to be a PB, especially when the pulse duration is less than 4μs. It should be pointed out that as the T increases, the α_{PB} gradually decreases to around 0.4. According to the results, overlaps exist in the parameter distribution of PB and NB, which lead to the difficulties to set an adequate threshold for multi-parameter classification. However, the turning points of α_{PB} -P and α_{PB} -T is generally consistent with the actual peak ratio and pulse duration distribution of PB and NB. This indicates the CAM values from the proposed model is helpful in threshold determination.

3.2.4 Narrow Bipolar events (NBs)

Narrow bipolar event (NB) is a special type of intracloud discharge, often occurring in isolation from other discharge events. The amplitude of NB is usually high, which can be close to that of RS(Rakov & Uman, 2003; Smith et al., 1999). The pulse duration of NB is short, ranging from 2 to 20μs (Jacobson and Light 2012; Wu et al. 2014). Increasing evidences indicate that NB may be the initiating process for other lightning discharge events. It is demonstrated through simulation that the electric field waveform of the NBE is related to the abrupted elongation of the initial negative leader channel in the thundercloud(Da Silva & Pasko, 2015). Because of the distribution overlap of several essential characteristic like the amplitude and pulse width between NB and RS, the NB is thus an important factor affecting the accuracy of RS classification in LLS.(Leal et al., 2019; Nag et al., 2014). In this section, the visualization of which part of the NB waveform causes the model to make the correct classification will be analyzed.

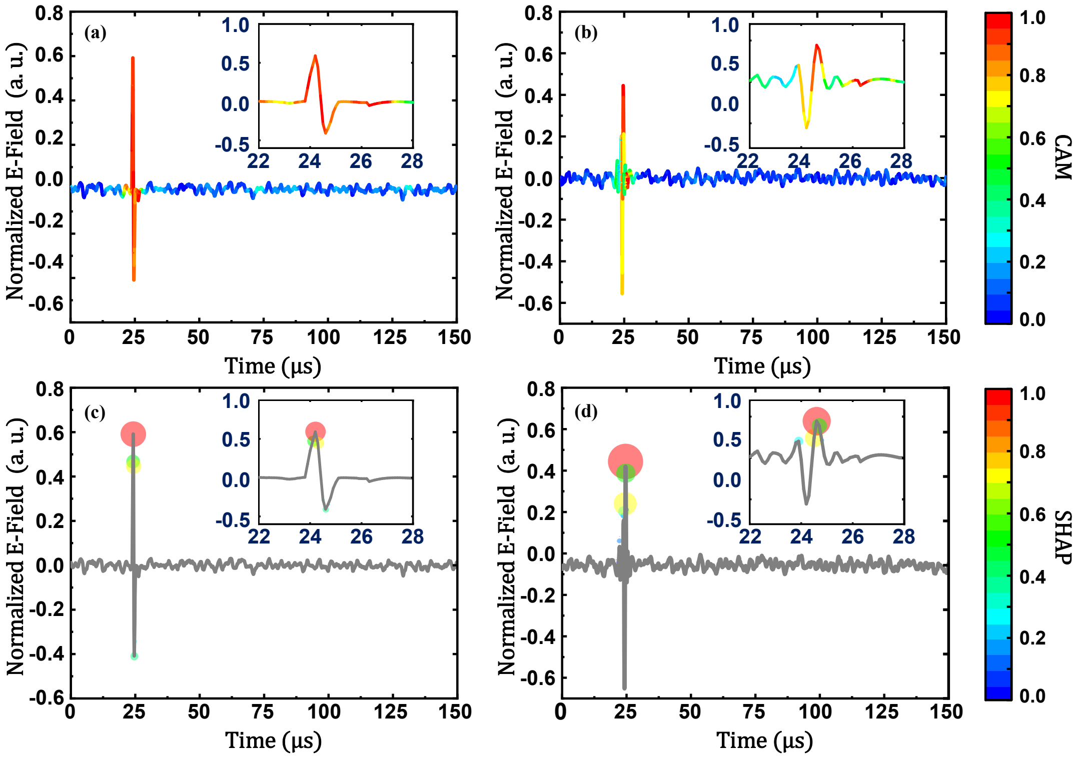


Figure 7 Visualized classification results of NB waveform #190818114306-001NB and #190818114919-001NB. **a)** Visualized CNN classification result based on CAM for the first NB case with a detailed demonstration for the main pulse part **b)** Visualized CNN classification result based on CAM for the second NB case with a detailed demonstration for the main pulse part **c)** Visualized SVM classification result based on SHAP for the first case and detailed demonstration as (a) **d)** Visualized SVM classification result based on SHAP for the second case and detailed demonstration as (b)

Figure 7(a) shows the visualized CNN classification result for an NB waveform, which is captured at 11:43:06 18th August 2019 at Anqing, Anhui, China. It can be seen that the proposed CNN locates the main pulse part of this waveform within 24.5±1.5μs accurately, which contains both positive and negative peaks as well as a steep polarity change process. It is notable that the average CAM value of the main pulse is higher than 0.8, while the average CAM value of the other part is less than 0.2, indicating that the CNN model focuses on the waveform pulse and pays less attention to the background waveform. The CAM value difference shows that the CNN model captures the feature of isolation of the NB waveform, with attention focused on the pulse part in the waveform which contains the most information of the waveform. Figure 7(b) shows the visualized SVM classification result for the same case. Compared to the CNN model, the SVM model also focuses on the pulse part and the high weight point is distributed at the positive and negative peak tops, with almost no high weight point distributed in the background part. Figure 7(c) shows that the weight values positive is positively correlated to the pulse magnitude, indicating that the SVM model may classify waveforms by identifying the high amplitude parts in the waveform. It is important to point out that the two core characteristics of the NB require the model to be able to recognize time-dependent feature. The short pulse duration feature requires the model to be able to determine if the pulse presents for a certain period of time. The bipolar pulse feature requires the model to be able to determine if positive and negative peaks appear in succession. As depicted in 2.3.2, the parallel design of multiple convolutional kernels in our CNN model allows the model to capture features at different time scales during the training process. Due to the variable scale of the features extracted by the CNN model, the model can learn the temporal causality features in waveforms during the learning process. Therefore, the hotspot region given out includes the entire main pulse part. In contrast, the SVM model only focuses on the high amplitude point distribution of the waveform and therefore may lead to misclassification of some NB waveforms. Figure 7(b) and Figure 7(d) show another case of NB. The bipolar pulse in this case is located around 23.9 to 24.9μs, but there is a large positive disturbance at 23.5μs with a peak ratio of approximately 0.5 to the main pulse. Figure 7(b) shows that the CNN model successfully classifies the waveform, with the given hotspot region distributed between 23.8-24.9μs, which is coincides with the main pulse duration. It should be noted that the CAM value of the positive part disturbance is less than 0.5. However, the SVM model is unable to classify this case correctly. As can be seen from Figure 7(d), the SVM model only locates the positive peak top of the main pulse as well as the positive peak top of the disturbance, ignoring the negative period of the main pulse, and thus makes incorrect judgments as a result. As can be seen from the above comparison, the SVM model only classifies waveforms by the distribution of high amplitude points, which lacks consideration of temporal correlation. The classification of SVM is not supported by physical process and is prone to misclassification. The CNN model takes into account the temporal correlation patterns in the timeseries data at an adaptive scale during the classification process, and is more likely to capture the core features of the NB waveform, providing higher classification accuracy and robustness.

To further demonstrate that the CNN model can help to find threshold to distinguish similar waveforms like RS and NB, we calculated the peak ratio P, pulse duration T and average CAM values α_{NB} of the NB and RS waveforms in the dataset. The definitions of P and T are identical to those in 3.2.3. The α_{NB} refers to the CAM value given by the model when the waveform is considered to be a NB. The result is shown in Figure 8.

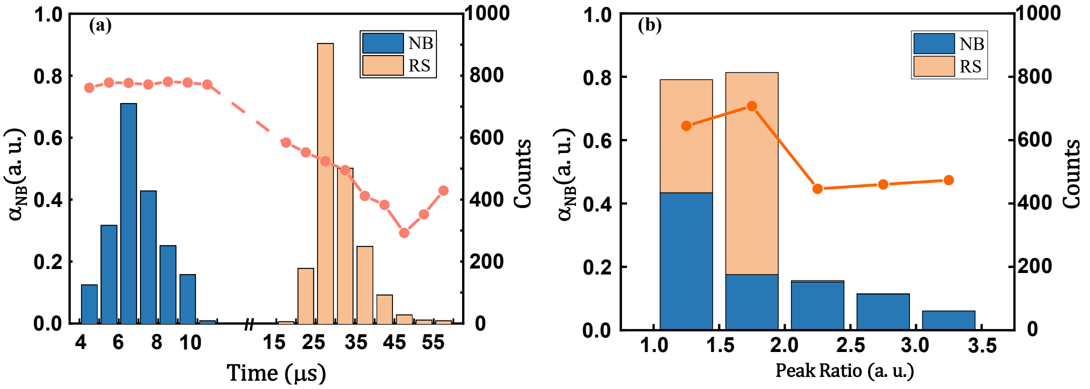


Figure 8 (a) relationship of CAM values α_{NB} and pulse duration T (b) relationship of CAM values α_{NB} and peak ratio P

Figure 8(a) shows the relationship between the average CAM value α_{NB} and the pulse duration T. When the pulse duration T is between 2-5.5μs, the average CAM value α_{NB} remains high within 0.75-0.8. As the T increases, α_{NB} drops rapidly to a minimum value of 0.22. It can be concluded from figure 8(a) that the model suggests that the pulse duration of the NB should not exceed 15μs. Figure 8(b) shows the relationship between the average CAM value α_{NB} and the peak ratio P. The peak ratio quantifies the degree of bipolarity of the waveform. When the P is between 1 and 2, the bipolarity of the waveform is more obvious and the average CAM value is at a high level between 0.6 and 0.8. When the P is greater than 2 and increases further, the bipolar characteristics of the waveform can be considered to have gradually disappeared and the unipolar characteristics become prominent. The CAM value drops steeply and remains at a low level of around 0.4. Figure 8(a) and (b) show that the turning points of α_{NB}-P and α_{NB}

-T is generally consistent with the actual peak ratio and pulse duration distribution of RS and NB.

4 Discussion

Compared to plain CNN models, our model employs the shortcut connection to improve the rate of convergence, which enables the model to deal with longer input waveforms. The multi-size kernels of our model can capture multi-scale temporal features of VLF/LF waveforms. In section 3.2, we demonstrate that the improved model can extract physical information of VLF/LF waveforms which is related to different lightning discharge processes. It means that the classification accuracy of our model is much less dependent on the dataset, and it can be extended to recognize the lightning VLF/LF waveform recorded from other regions.

The open dataset measured at Córdoba in central Argentina (Zhu et al., 2021) is employed as a test dataset here to further exam the portability of our model. The waveforms in the original dataset are classified into four types, including +CG, - CG, + IC and – IC. We ignore the impact of polarity on the classification accuracy since it is convenient to be recognized. The original dataset is reclassified into two groups of RS and IC. We down sample these waveforms to the sample rate of 5MS/s, which is corresponding to our devices. Since the input waveform of our model must possess 2500 points, we manually add noise data to the waveforms meet the required length. Higher classification accuracy is obtained by using our model than the result reported by Zhu et al. The classification accuracy of RS in our model is 99.41%, while it is 97% by using SVM. The classification accuracy of IC in our model is 97.38%, while it is 97% by using SVM. Note that the classification accuracy of IC here equals to the sum of PB, NB, and ASIC for convenience of comparison.

It should be emphasized that traditional machine learning methods can only give out the probability for different categories. The proposed model can visualize the contribution of different parts of the waveform to the classification result. One can observe in figure 9 that our model can effectively capture the physical features of waveforms which cannot be classified correctly by the SVM. Fig. 9 (a) shows the CAM visualization of a RS event. This case is misclassified as an IC by the SVM, which may be caused by the unexpected bipolar oscillation around 30μs according to the discussion of Zhu et al. It can be seen from figure 9(a) that the proposed model marks the dual peaks of the waveform which is believed to be an instinct feature of RS events. Moreover, the unexpected bipolar oscillation is neglected by our model, since the corresponding CAM are less than 0.5 there. Figure 9(b) shows the CAM visualization of a NB event. It is misclassified as a CG by the SVM for the positive disturbance around 27 μs according to Zhu's report. It can be seen in figure 9(b) that the steep polarity change process is marked with high CAM values, which is the key feature in the NB waveforms. The positive disturbance is neglected with the CAM values less than 0.5. According to these comparisons, the proposed CNN model is more effective in extracting the key features related to the physical process. These key features are relatively invariant in different regions. Therefore, the CNN model is able to accurately classify data from different regions.

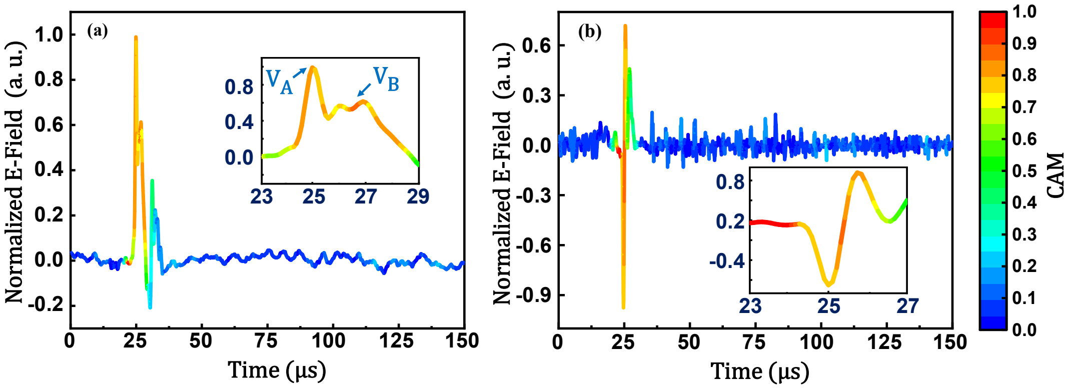


Figure 9 The CAM values of two misclassified waveforms by the SVM in the open dataset. (a) a RS waveform misclassified as IC. (b) an IC waveform misclassified as RS

5 Conclusion

In this paper, the main conclusions are summarized as follows:

（1）In this paper, an interpretable CNN model for VLF/LF lightning waveform classification is proposed. The proposed model uses multi-scale convolutional kernels to enhance the ability to capture local waveform features. The output of the final convolutional layer and the fully-connection weights are used to visualize the contribution of different waveform parts to the classification result. A shortcut connection is built in the proposed CNN model to promote the convergence speed and make the model capable of waveforms with higher sampling rate. Based on 8000 waveforms recorded in five provinces in China, the four-type classification of waveforms including RS, ASIC, PB and NB, is achieved with an accuracy of 98.5%, which is better than the traditional SVM and RF methods.

（2）Based on the distribution of the high-contribution waveform parts in the classification process, we analyzed the correlation between the model's focused features and the lightning discharge process. The model considers the double peak structure superimposed on the main pulse as the main feature of the RS, which is mainly caused by the abrupt change of the current or the branches of the lightning channel. The proposed model is able to identify the separated, repetitive pulses generated by the ASIC event, which are associated with the stepped growth of the negative leader in thunderstorms. The continuous bipolar pulse train is considered as the main feature of PB by the model, which is generated by the continuous development of the initial leader in thunderstorms. The model in this paper is also able to identify the narrow bipolar pulse generated by the sudden elongation of the initial leader in NB events. This indicates that compared to traditional machine learning methods, the model in this paper extracts features which are in line with the human experts in the VLF/VF waveform model classification process.

（3）We analyzed the relationship between the average waveform weights given by the model and the pulse duration, peak ratios. For NB and PB events with similar physical mechanisms, the weight value α_{PB} for PB events varies in a U-shape with the increase of the peak ratio. When the pulse duration T_{width} is greater than 4.0 μs, α_{PB} decreases monotonically with the increase of T_{width}. This indicates that the pulse duration is more suitable than the peak ratio to distinguish the NB and PB waveforms. Compared to the RS, the weight value α_{NB} for NB does not vary significantly (which is between 0.4 and 0.7) with the peak ratio. And when the pulse duration T_{width} is greater than 5.0 μs, α_{NB} decreases significantly with the increase of T_{width}, which indicates that the pulse duration can better solve the problem of easy confusion between RS and NB events in the lightning location system.

（4）We validated the model in this paper using an open source dataset reported in literature, which has a total of 32,754 samples from central Argentina. The model in this paper achieved an accuracy of 98.39%, which is better than the result using the SVM according to the literature. Based on the contribution weights obtained in this paper, it can be seen that the model in this paper considers the double-peaked structure superimposed on the main pulse as the key feature of the RS, which avoids the influence of unexpected waveform oscillation and NB waveforms on the RS/IC classification accuracy. It is proved that the model in this paper not only reduces the dependence of the classification performance on the training set, but also is more robust in the classification of waveforms from different regions.

Acknowledgement

This work is supported by Science and Technology Project of State Grid Corporation of China (5500-2021120583A-0-5-SF).

Open Research

The interpretable CNN model proposed in this study and related data(Xiao et al., 2022) are available at: https://doi.org/10.5281/zenodo.7549481

References

Biagi, C. J., Cummins, K. L., Kehoe, K. E., & Krider, E. P. (2007). National lightning detection network (NLDN) performance in southern Arizona, Texas, and Oklahoma in 2003–2004. *Journal of Geophysical Research: Atmospheres*, *112*(D5). https://doi.org/10.1029/2006JD007341

Bils, J. R., Thomson, E. M., Uman, M. A., & Mackerras, D. (1988). Electric field pulses in close lightning cloud flashes. *Journal of Geophysical Research: Atmospheres*, *93*(D12), 15933–15940. https://doi.org/10.1029/JD093iD12p15933

Bitzer, P. M., Christian, H. J., Stewart, M., Burchfield, J., Podgorny, S., Corredor, D., Hall, J., Kuznetsov, E., & Franklin, V. (2013). Characterization and applications of VLF/LF source locations from lightning using the Huntsville Alabama Marx Meter Array. *Journal of Geophysical Research: Atmospheres*, *118*(8), 3120–3138. https://doi.org/10.1002/jgrd.50271

Brunner, K. N. (2016). *Explorations in intracloud lightning and leader processes*. The University of Alabama in Huntsville.

Cooray, V. (2009). Propagation Effects Due to Finitely Conducting Ground on Lightning-Generated Magnetic Fields Evaluated Using Sommerfeld's Integrals. *IEEE Transactions on Electromagnetic Compatibility*, *51*(3), 526–531. https://doi.org/10.1109/TEMC.2009.2019759

Cooray, V., & Lundquist, S. (1982). On the characteristics of some radiation fields from lightning and their possible origin in positive ground flashes. *Journal of Geophysical Research: Oceans*, *87*(C13), 11203–11214. https://doi.org/10.1029/JC087iC13p11203

Cooray, V., & Lundquist, S. (1985). Characteristics of the radiation fields from lightning in Sri Lanka in the tropics. *Journal of Geophysical Research: Atmospheres*, *90*(D4), 6099–6109. https://doi.org/10.1029/JD090iD04p06099

Da Silva, C. L., & Pasko, V. P. (2015). Physical mechanism of initial breakdown pulses and narrow bipolar events in lightning discharges. *Journal of Geophysical Research: Atmospheres*, *120*(10), 4989–5009. https://doi.org/10.1002/2015JD023209

Gomes, C., & Cooray, V. (2001). Characteristics of cloud flashes. *14th International Zurich Symposium EMC 58J3*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. https://doi.org/10.1109/cvpr.2016.90

Kašpar, P., Santolík, O., Kolmašová, I., & Farges, T. (2017). A model of preliminary breakdown pulse peak currents and their relation to the observed electric field pulses. *Geophysical Research Letters*, *44*(1), 596–603. https://doi.org/10.1002/2016gl071483

Kaur, H., Pannu, H. S., & Malhi, A. K. (2019). A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions. *ACM Comput. Surv.*, *52*(4). https://doi.org/10.1145/3343440

Kohlmann, H., Schulz, W., & Pedeboy, S. (2017). Evaluation of EUCLID IC/CG classification performance based on ground-truth data. *2017 International Symposium on Lightning Protection (XIV SIPDA)*, 35–41. https://doi.org/10.1109/SIPDA.2017.8116896

Krider, E. P., Radda, G. J., & Noggle, R. C. (1975). Regular radiation field pulses produced by intracloud lightning discharges. *Journal of Geophysical Research*, *80*(27), 3801–3804. https://doi.org/10.1029/jc080i027p03801

Le Vine, D. M. (1980). Sources of the strongest RF radiation from lightning. *Journal* *of Geophysical Research: Oceans*, *85*(C7), 4091–4095. https://doi.org/10.1029/jc085ic07p04091

Leal, A. F., Rakov, V. A., & Rocha, B. R. (2019). Compact intracloud discharges: New classification of field waveforms and identification by lightning locating systems. *Electric Power Systems Research*, *173*, 251–262. https://doi.org/10.1016/j.epsr.2019.04.016

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57.

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Lyu, F., Cummer, S. A., & McTague, L. (2015). Insights into high peak current in-cloud lightning events during thunderstorms. *Geophysical Research Letters*, *42*(16), 6836–6843. https://doi.org/10.1002/2015GL065047

Mallick, S., Rakov, V. A., Hill, J. D., Ngin, T., Gamerota, W. R., Pilkey, J. T., Jordan, D. M., Uman, M. A., Heckman, S., Sloop, C. D., & Liu, C. (2015). Performance characteristics of the ENTLN evaluated using rocket-triggered lightning data. *Electric Power Systems Research*, *118*, 15–28. https://doi.org/10.1016/j.epsr.2014.06.007

Murphy, M. J., Cramer, J. A., & Said, R. K. (2021). Recent History of Upgrades to the U.S. National Lightning Detection Network. *Journal of Atmospheric and Oceanic Technology*, *38*(3), 573–585. https://doi.org/10.1175/JTECH-D-19-

0215.1

Nag, A., Murphy, M. J., Cummins, K. L., Pifer, A. E., & Cramer, J. A. (2014). Recent evolution of the us National lightning detection network. *23rd International Lightning Detection Conference & 5th International Lightning Meteorology Conference*.

Nag, A., & Rakov, V. A. (2008). Pulse trains that are characteristic of preliminary breakdown in cloud-to-ground lightning but are not followed by return stroke pulses. *Journal of Geophysical Research: Atmospheres*, *113*(D1). https://doi.org/10.1029/2007JD008489

Nassralla, M., El Zein, Z., & Hajj, H. (2017). Classification of normal and abnormal heart sounds. *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)*, 1–4. https://doi.org/10.1109/ICABME.2017.8167538

Peng, C., Liu, F., Zhu, B., & Wang, W. (2019). A convolutional neural network for classification of lightning LF/VLF waveform. *2019 11th Asia-Pacific International Conference on Lightning (APL)*, 1–4. https://doi.org/10.1109/APL.2019.8815977

Rakov, V. A., & Uman, M. A. (2003). *Lightning: Physics and effects*. Cambridge university press.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “ Why should i trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Said, R., Inan, U., & Cummins, K. (2010). Long-range lightning geolocation using a VLF radio atmospheric waveform bank. *Journal of Geophysical Research: Atmospheres*, *115*(D23). https://doi.org/10.1029/2010JD013863

Schulz, W., Diendorfer, G., Pedeboy, S., & Poelman, D. R. (2016). The European lightning location system EUCLID –\hack\newline Part 1: Performance analysis and validation. *Natural Hazards and Earth System Sciences*, *16*(2), 595–605. https://doi.org/10.5194/nhess-16-595-2016

Shao, X.-M., & Jacobson, A. R. (2009). Model Simulation of Very Low-Frequency and Low-Frequency Lightning Signal Propagation Over Intermediate Ranges. *IEEE Transactions on Electromagnetic Compatibility*, *51*(3), 519–525. https://doi.org/10.1109/TEMC.2009.2022171

Smith, D., Shao, X., Holden, D., Rhodes, C., Brook, M., Krehbiel, P., Stanley, M., Rison, W., & Thomas, R. (1999). A distinct class of isolated intracloud lightning discharges and their associated radio emissions. *Journal of Geophysical Research: Atmospheres*, *104*(D4), 4189–4212. https://doi.org/10.1029/1998JD200045

Wang, J., Huang, Q., Ma, Q., Chang, S., He, J., Wang, H., Zhou, X., Xiao, F., & Gao, C. (2020). Classification of VLF/LF lightning signals using sensors and deep learning methods. *Sensors*, *20*(4), 1030. https://doi.org/10.3390/s20041030

Wang, Y., Gu, S., Fang, Y., Xu, Y., Chen, Y., & Li, P. (2020). Compact electric field change meter and its application in lightning detection and fault analysis for power grids. *2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, 1–4. https://doi.org/10.1109/ICHVE49031.2020.9279853

Wang, Y., Qie, X., Wang, D., Liu, M., Su, D., Wang, Z., Liu, D., Wu, Z., Sun, Z., & Tian, Y. (2016). Beijing Lightning Network (BLNET) and the observation on preliminary breakdown processes. *Atmospheric Research*, *171*, 121–132. https://doi.org/10.1016/j.atmosres.2015.12.012

Wooi, C.-L., Abdul-Malek, Z., Salimi, B., Ahmad, N. A., Mehranzamir, K., & Vahabi-Mashak, S. (2015). A comparative study on the positive lightning return stroke electric fields in different meteorological conditions. *Advances in Meteorology*, *2015*. https://doi.org/10.1155/2015/307424

Wu, T., Wang, D., & Takagi, N. (2018). Locating preliminary breakdown pulses in positive cloud-to-ground lightning. *Journal of Geophysical Research: Atmospheres*, *123*(15), 7989–7998. https://doi.org/10.1029/2018JD028716

Xiao, L., Wang, Y., He, H., & Chen, W. (2022). VLF/LF lightning waveform classification version-alpha[dataset]. https://doi.org/10.5281/zenodo.7549481

Zhu, Y., Bitzer, P., Rakov, V., & Ding, Z. (2021). A machine-learning approach to classify cloud-to-ground and intracloud lightning. *Geophysical Research Letters*, *48*(1), e2020GL091148. https://doi.org/10.1029/2020GL091148

Figure1.

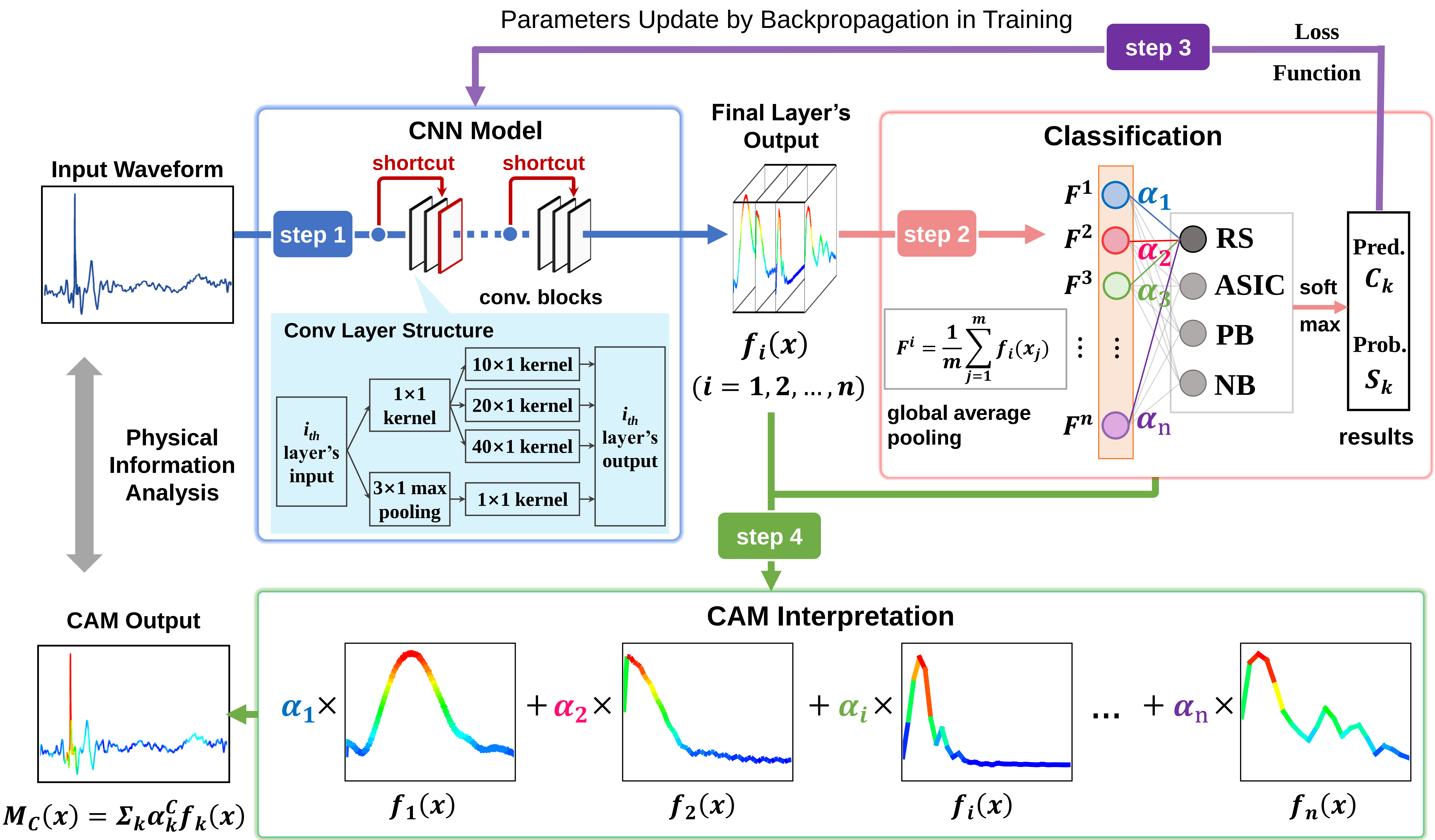

Parameters Update by Backpropagation in Training
step 3
Loss Function
Input Waveform
CNN Model
shortcut
shortcut
step 1
conv. blocks
Conv Layer Structure
i_{th} layer's input
1×1 kernel
3×1 max pooling
10×1 kernel
20×1 kernel
40×1 kernel
1×1 kernel
i_{th} layer's output
Final Layer's Output
$f_i(x)$
$(i = 1, 2, \ldots, n)$
Classification
step 2
$F^i = \frac{1}{m}\sum_{j=1}^{m} f_i(x_j)$
global average pooling
F^1
F^2
F^3
F^n
α_1
α_2
α_3
α_n
RS
ASIC
PB
NB
soft max
Pred. C_k
Prob. S_k
results
Physical Information Analysis
step 4
CAM Output
$M_C(x) = \Sigma_k \alpha_k^C f_k(x)$
CAM Interpretation
$\alpha_1 \times$
$+ \alpha_2 \times$
$+ \alpha_i \times$
…
$+ \alpha_n \times$
$f_1(x)$
$f_2(x)$
$f_i(x)$
$f_n(x)$

Figure2.

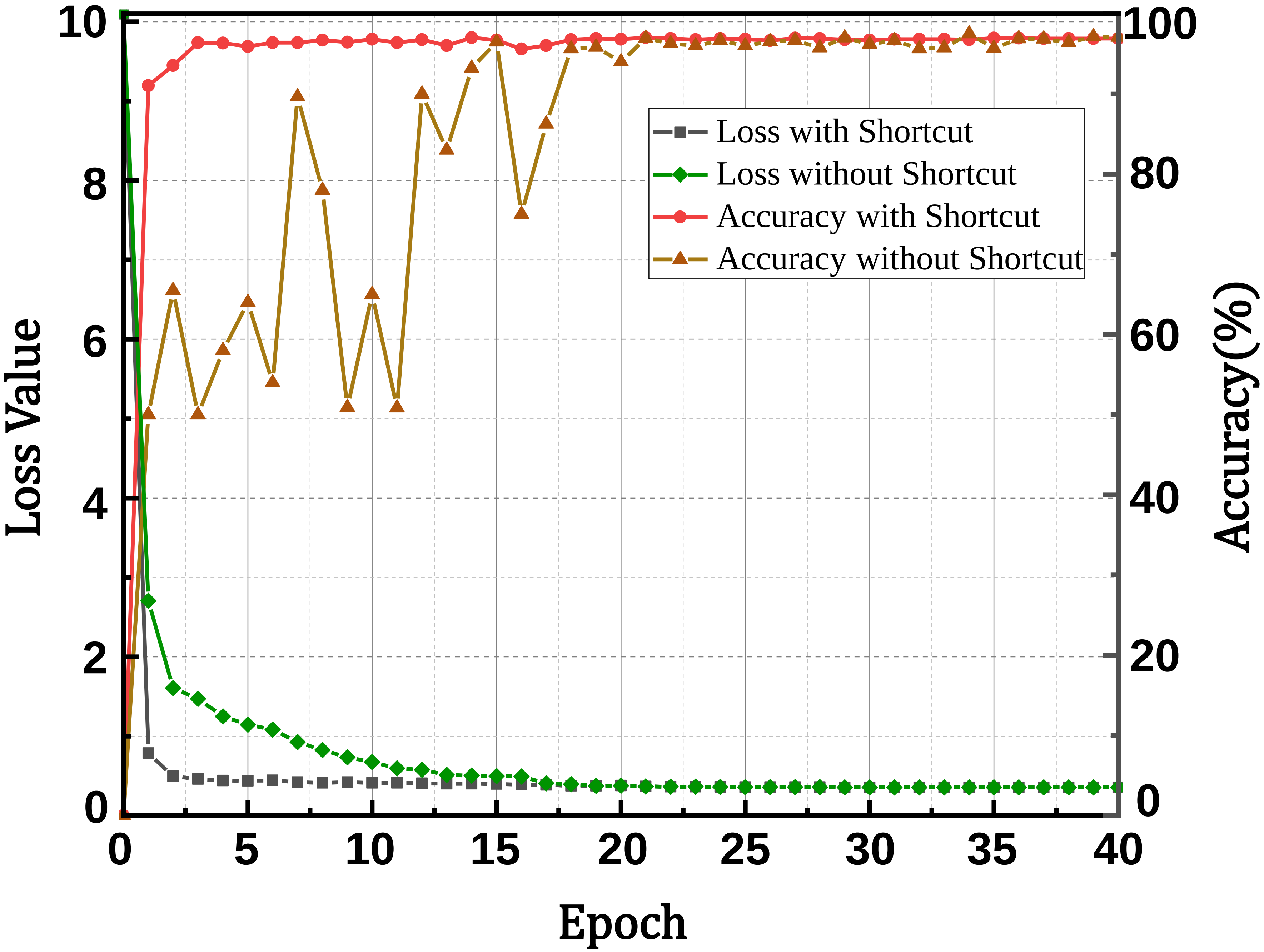

Loss with Shortcut
Loss without Shortcut
Accuracy with Shortcut
Accuracy without Shortcut
Loss Value
Accuracy(%)
Epoch
0
2
4
6
8
10
0
20
40
60
80
100
0
5
10
15
20
25
30
35
40

Figure3.

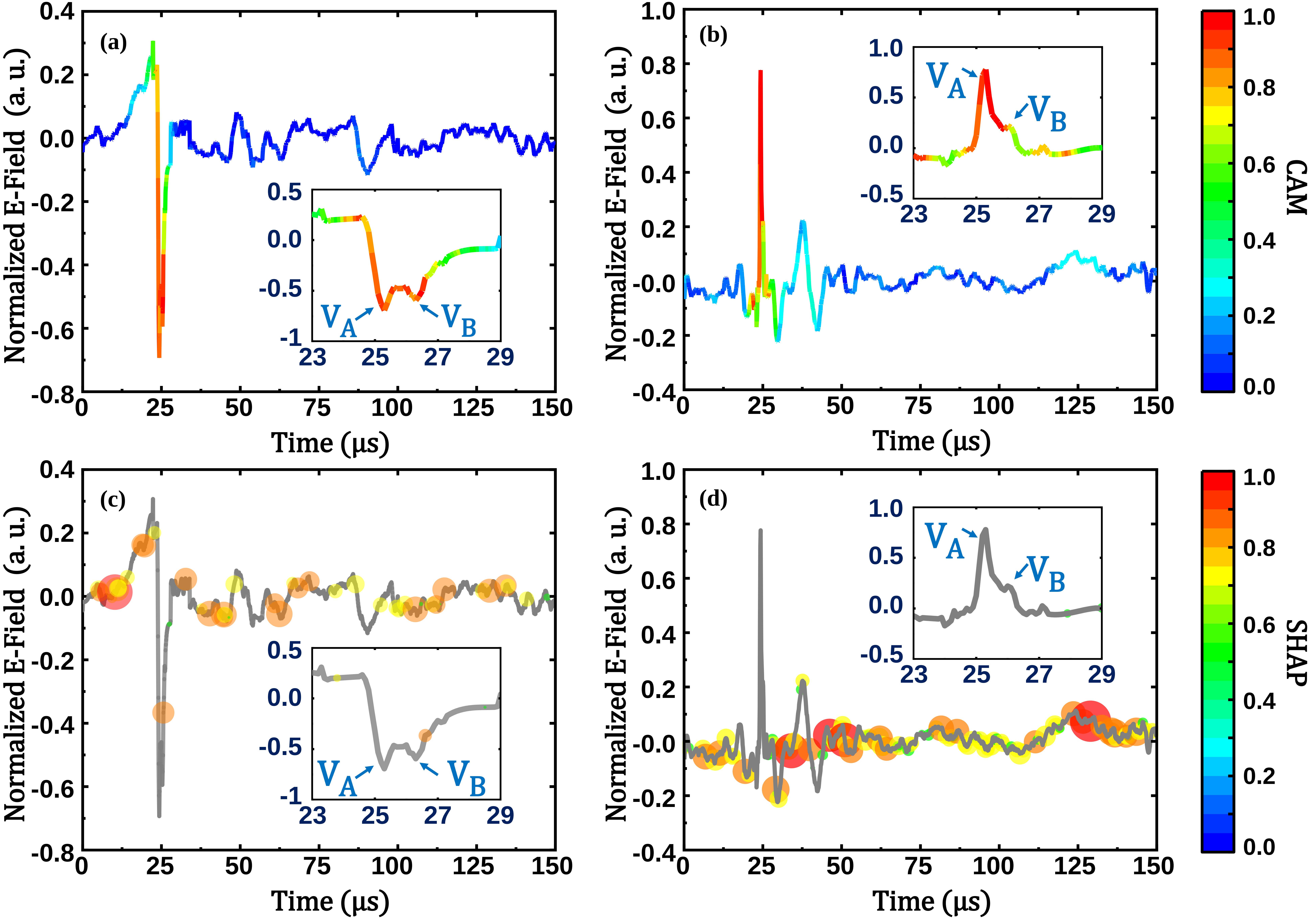

(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (µs)
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
1.0
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
0
25
50
75
100
125
150
0.5
-0.5
-1
23
25
27
29
V_A
V_B
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0

Figure4.

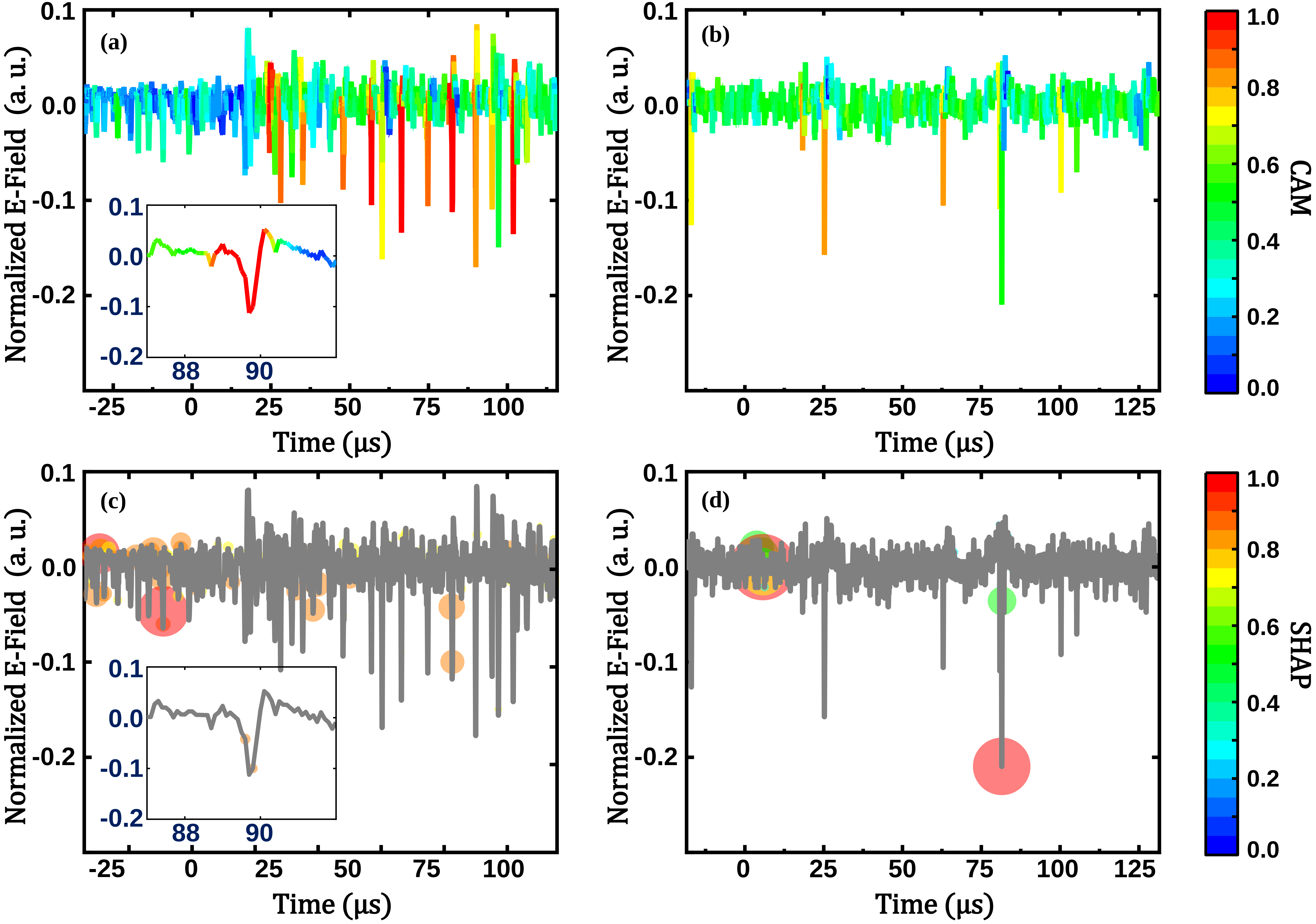

(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (μs)
0.1
0.0
-0.1
-0.2
-25
0
25
50
75
100
125
88
90
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0

Figure5.

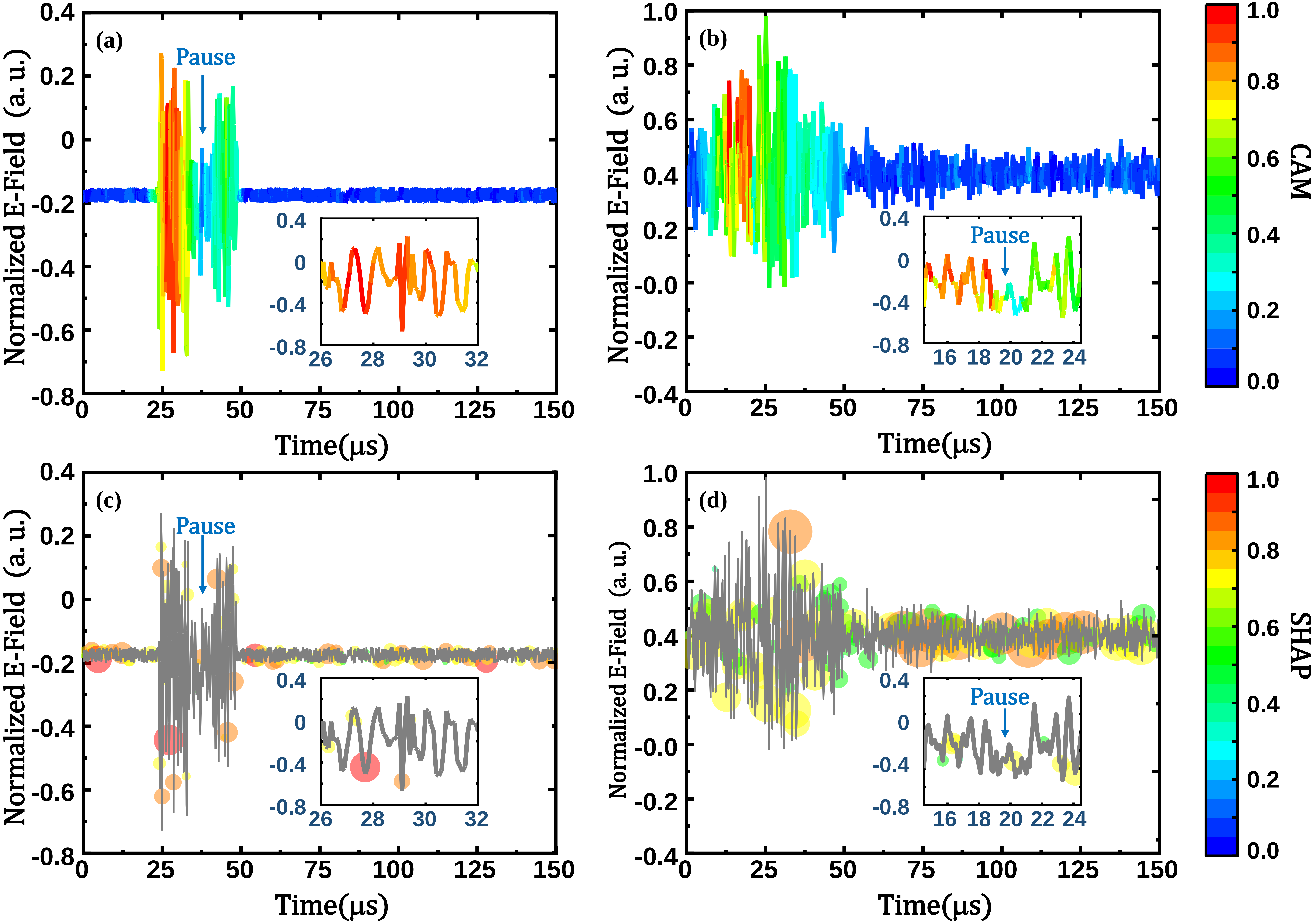

(a)
(b)
(c)
(d)
Pause
Normalized E-Field (a. u.)
Time(μs)
0
25
50
75
100
125
150
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1.0
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
26
28
30
32
16
18
20
22
24
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0

Figure6.

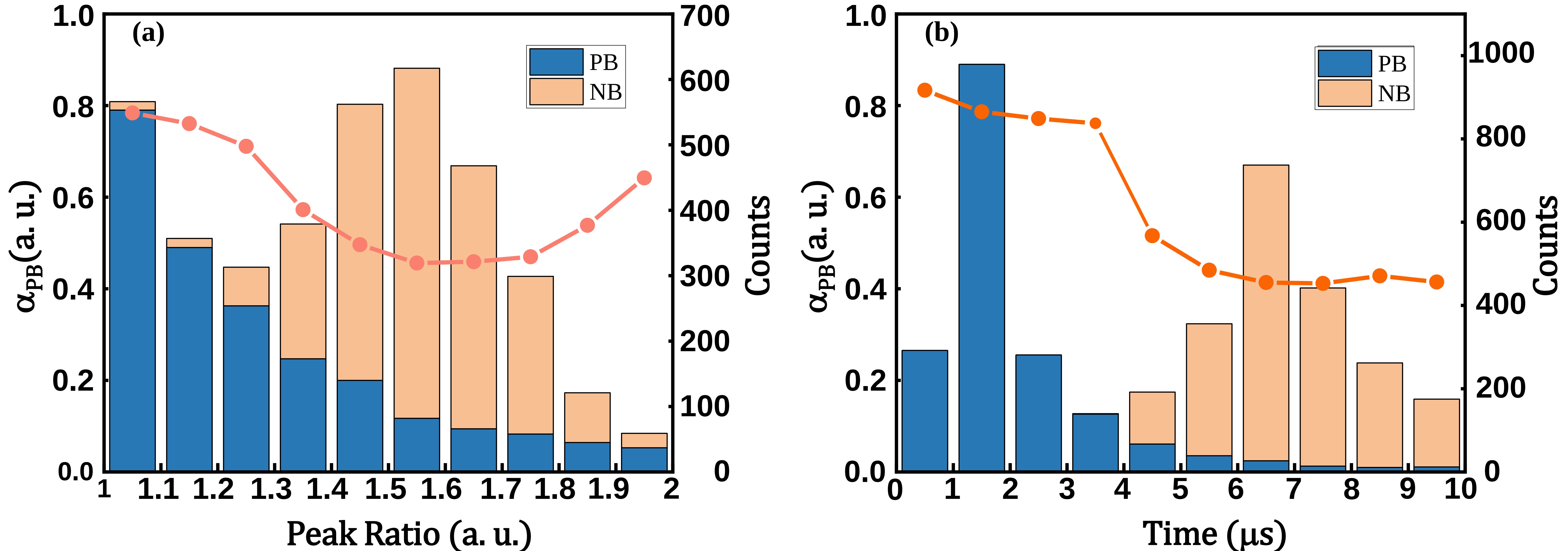

(a)
PB
NB
α_{PB}(a. u.)
1.0
0.8
0.6
0.4
0.2
0.0
Counts
700
600
500
400
300
200
100
0
1
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2
Peak Ratio (a. u.)
(b)
PB
NB
α_{PB}(a. u.)
1.0
0.8
0.6
0.4
0.2
0.0
Counts
1000
800
600
400
200
0
0
1
2
3
4
5
6
7
8
9
10
Time (μs)

Figure7.

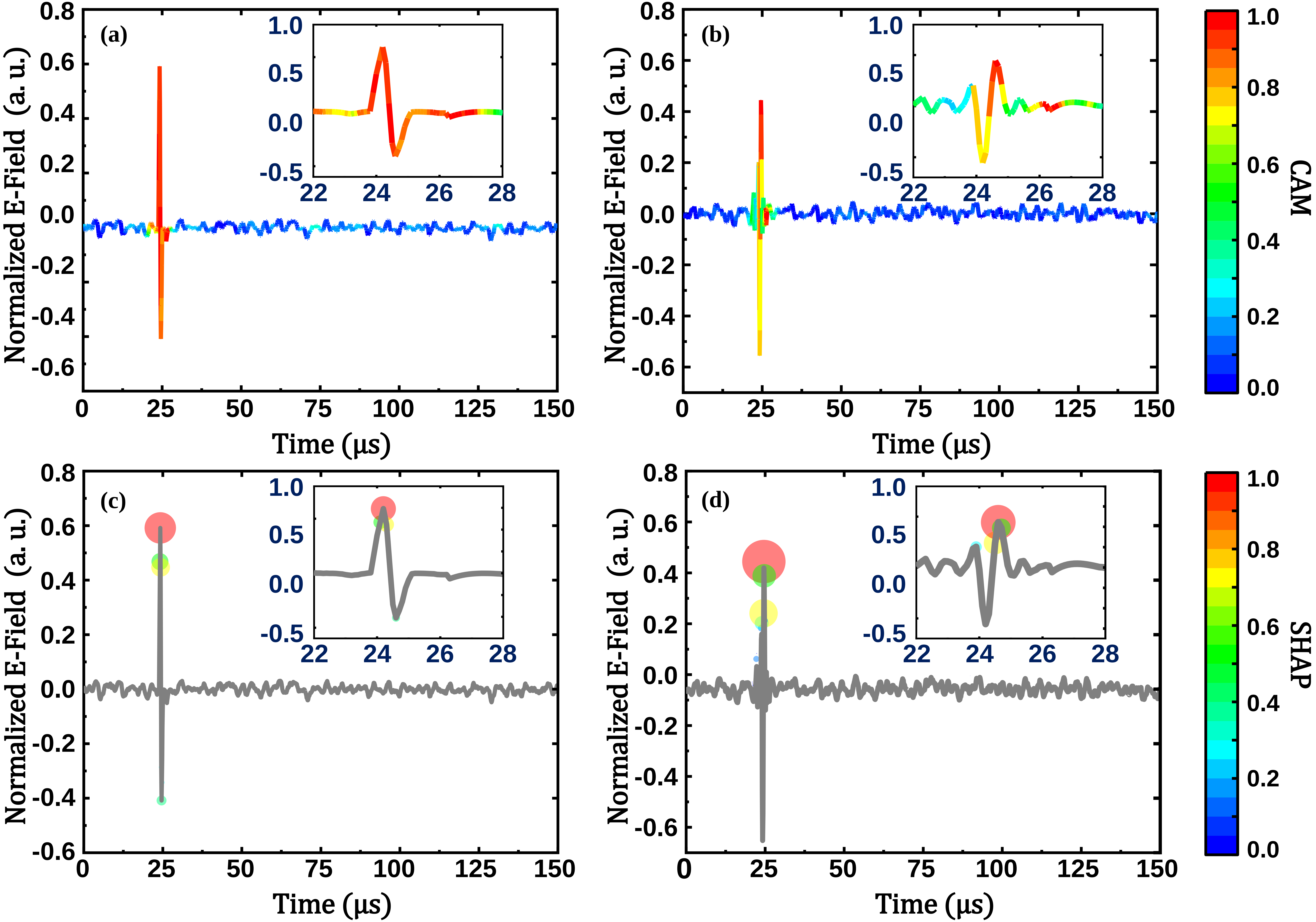

(a)
(b)
(c)
(d)
Normalized E-Field (a. u.)
Time (μs)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
25
50
75
100
125
150
1.0
0.5
0.0
-0.5
22
24
26
28
CAM
SHAP
1.0
0.8
0.6
0.4
0.2
0.0

Figure8.

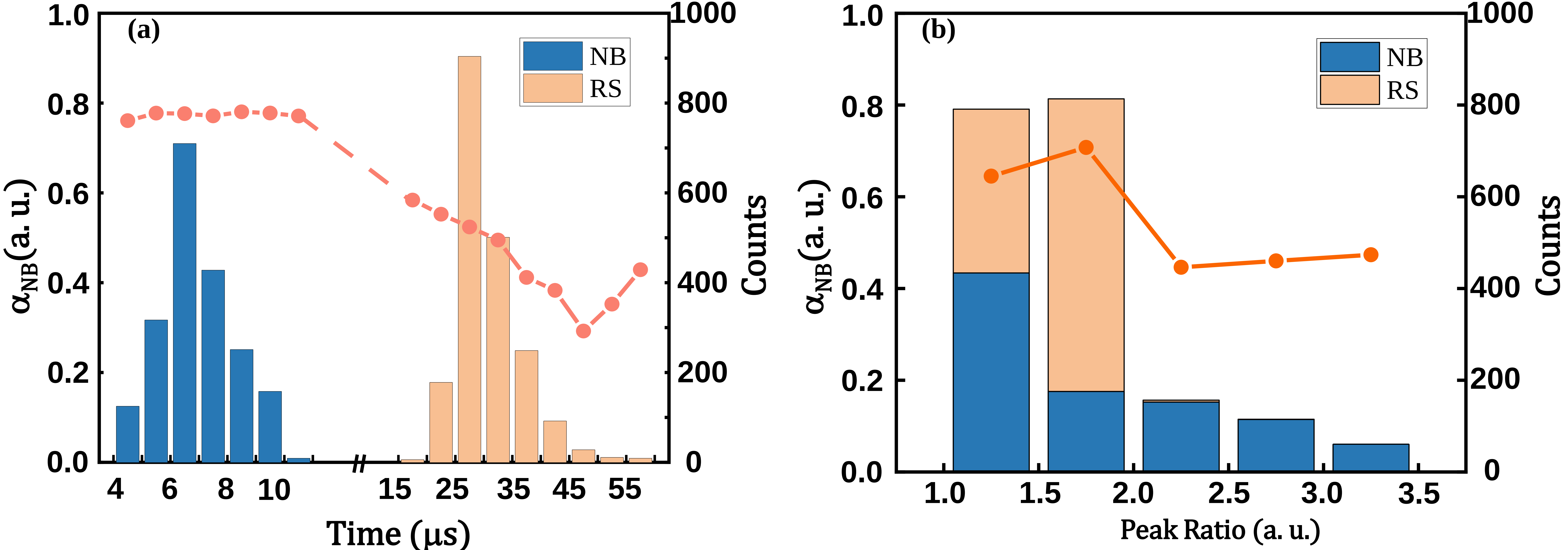
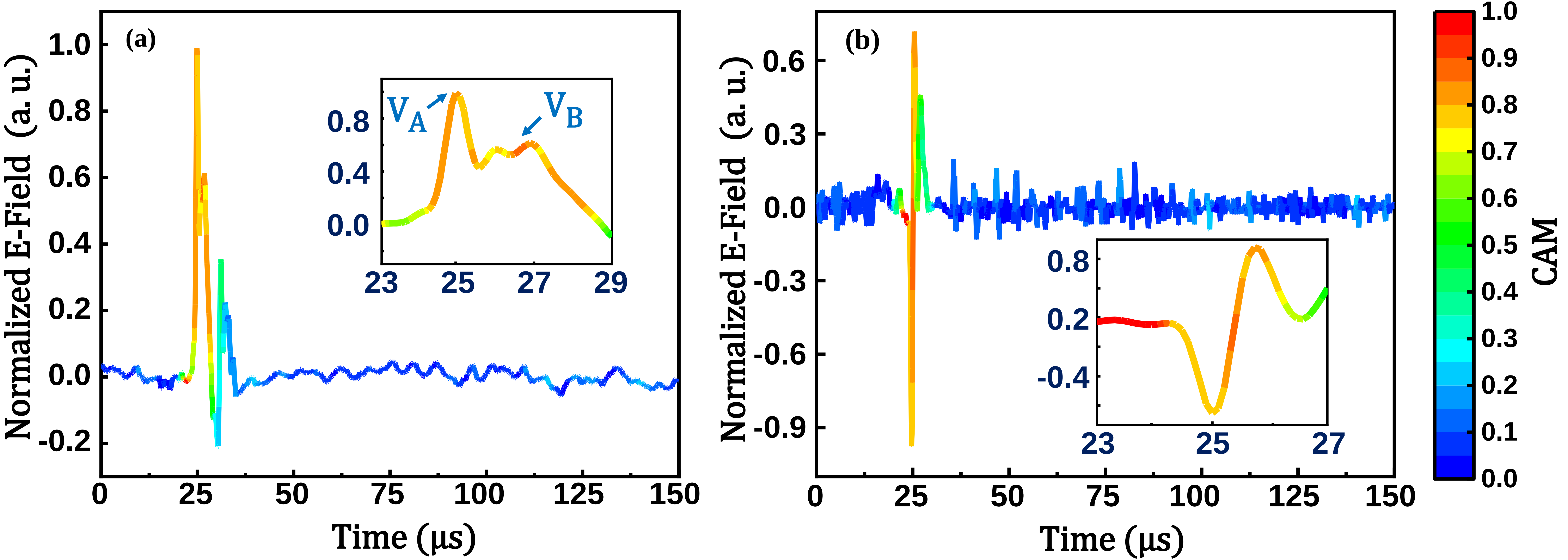

(a)
NB
RS
α_{NB}(a. u.)
1.0
0.8
0.6
0.4
0.2
0.0
1000
800
600
400
200
0
Counts
4
6
8
10
15
25
35
45
55
Time (μs)
(b)
NB
RS
α_{NB}(a. u.)
1.0
0.8
0.6
0.4
0.2
0.0
1000
800
600
400
200
0
Counts
1.0
1.5
2.0
2.5
3.0
3.5
Peak Ratio (a. u.)

Figure9.


(a)
Normalized E-Field (a. u.)
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
25
50
75
100
125
150
Time (μs)
V_A
V_B
0.8
0.4
0.0
23
25
27
29
(b)
Normalized E-Field (a. u.)
0.6
0.3
0.0
-0.3
-0.6
-0.9
0
25
50
75
100
125
150
Time (μs)
0.8
0.2
-0.4
23
25
27
CAM
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Towards an Interpretable CNN Model for the Classification of Lightning Produced VLF/LF Signals

Lilang Xiao[1], Weijiang Chen[4], Yu Wang[3], Kai Bian[4], Zhong Fu[2], Nianwen Xiang[5], Hengxin He[1], Yang Cheng[2]

[1] State Key Laboratory of Advanced Electromagnetic Engineering and Technology, HUST, Wuhan, People's Republic of China

[2] Electric Power Research Institute, State Grid Anhui Electric Power Company, Hefei, People's Republic of China

[3] Wuhan NARI Co., Ltd of State Grid Electric Power Research Institute, People's Republic of China

[4] State Grid of China, Beijing, People's Republic of China

[5] Hefei University of Technology, School of Electrical Engineering and Automation, People's Republic of China

Corresponding author: Hengxin He (hengxin_he@hust.edu.cn)

Key Points:

- The proposed model can extract decisive features of VLF/LF lightning signals which is similar to the human expert's behavior.
- The model achieved an accuracy of 98.5% on a four-type lightning VLF/LF electrical waveforms dataset.
- Testing with data from Argentina validates that the accuracy of the model is less dependent on training data set.

Abstract:

Classification of lightning produced VLF/LF signals plays crucial role in the detection and location of lightning flashes. The machine learning method has potential in the VLF/LF lightning signal classification. Traditional machine learning methods are data-driven and work in a black-box fashion, making the classification accuracy highly dependent on the size and quality of dataset. In this paper, an interpretable convolutional neural network model is proposed for VLF/LF lightning electric field waveform classification. Multi-scale convolutional kernels and shortcut connections are adopted in this model to enhance the ability to capture local waveform features. The CAM method is embedded in our model to open the black-box by visualizing the weight of different waveform features on the classification results. Based on the measured data from five different provinces in China, an accuracy of 98.5% is achieved in a four-type classification task including RS, active stage of IC, PB and NB. The correlation between the weight values of different waveform features and corresponding lightning discharge process are analyzed. It is found that the proposed model can extract decisive features of VLF/LF lightning signals closely related to the physical process of lightning discharges, which is similar to the human expert's behavior. The proposed model is validated by using an open-source dataset from Argentina. It is indicated that the proposed model can resist the impact of unexpected waveform oscillation and achieve a higher accuracy of 98.39% than that of the support vector method. It is demonstrated that our model is less dependent on the training dataset.

Plain Language Summary

Electromagnetic waveforms in very low frequency and low frequency (VLF/LF) band are usually used to detect and locate different lightning activities. Traditional classification methods often misclassify in multi-type lightning discharge waveform classification. The machine learning models show promising potential in the multi-type classification task. However, these models cannot explain which part of the input waveform leads to the classification result, which makes the classification model unreliable. In this paper, we propose an improved and interpretable convolution neural network model, which is adapted to the lightning waveform classification task with changes in model structure. By utilizing the convolution outputs, the model can visualize the contribution of different parts of the waveform to the classification result. The analysis of the visualization results show that the high accuracy and generalization of the proposed model comes from the capture of waveform features corresponding to the key physical process in waveform generation. The dataset for model training comes from five provinces in China, which contains different meteorological conditions. The trained model based on the dataset reached a classification accuracy of 98.5% on test set and 98.39% on another open-source dataset from Argentina, which validated the generalization of the proposed model.

1 Introduction

Remote sensing the electromagnetic radiation generated by lightning discharges is an effective approach to detect and locate lightning activities. It is recognized that the radio emission in the VHF regime is primarily emitted by the streamer and leader involved in lightning discharges, while most of the radiation power is concentrated in the VLF/LF band that is mainly produced by the return stroke (RS) in cloud-to-ground flashes (CGs) and the active stage of intro-cloud flashes (ICs). The detection of VLF/LF radiation was initially introduced to sense the occurrence of CG remotely. Combined with the VLF/LF sensing and the time of arrival (TOA) method, the lightning location system (LLS) was proposed in 1980s which becomes an important technique to support the lightning protection for ground infrastructures nowadays. In order to exclude the impact of ICs, a fundamental task of LLS based on the VLF/LF detection is to recognize the characteristic waveforms produced by return strokes. In recent years, with the development of hardware performance, the emission source location of CGs and ICs can be achieved by using the short-baseline VLF/LF sensing technique and the 3D TOA method. The updated VLF/LF system not only can be utilized as an effective tool for lightning protection engineering applications, but also has the potential in lightning physics research. The lightning leader development process were investigated by using this technique, including the propagation of negative downward leader, the preliminary breakdown (PB), and the narrow bipolar event (NBE) etc. (Bitzer et al., 2013; Y. Wang et al., 2016; Wu et al., 2018). In order to improve the performance of the short-baseline system in lightning detection and lightning physics research, challenges arise in the accurate and automatic identification of waveform characteristics that produced by different lightning discharges.

For most LLS, the multi-parameter method is employed as the criterion to classify the CG and IC, which is derived from extensive field records(Murphy et al., 2021). It adopts specific parameters that can describe the primary profile of VLF/LF waveform, such as the amplitude, the rise and fall time, and the zero-cross time, etc. According to the results of validation studies, the RS detection efficiency of typical lightning location systems (including National Detection Networks (NLDN) and Earth Networks Total Lightning Network (ENTLN) in US, European Cooperation for Lightning Detection network (EUCLID) in Europe ranges from 71% to 92%, while the ICs detection efficiency varies from 73% to 96%(Biagi et al., 2007; Mallick et al., 2015; Schulz et al., 2016). Despite the difference in hardware performance, the deviation in detection efficiency of different systems is mainly attributed to the classification accuracy of CGs and ICs. On the one hand, since the multi-parameter method is difficult to extract characteristic parameters from VLF/LF signals with low-amplitude, the small signals were often abandoned, resulting in the decrease of detection efficiency(Kohlmann et al., 2017; Nag et al., 2014). On the other hand, the characteristic parameter involved in the multi-parameter method may vary in regions with different meteorological conditions(Cooray, 2009; Said et al., 2010; Shao & Jacobson, 2009; Wooi et al., 2015). For instance, the rise time and zero cross time of RS in Vitemölla, Sweden is of 5-25μs and approximately 40μs respectively, while the rise time decreases to about 2.5-9μs and the zero cross time increases to the range of 40-160μs in Sri Lanka(Cooray & Lundquist, 1982, 1985). Accurately determining the thresholds of characteristic parameters requires the support of long-term data. Recently, the machine learning methods such as the support vector machines (SVM) and the convolutional neural networks (CNN) are introduced to improve the classification efficiency of lightning VLF/LF signals. The SVM method is utilized to classify the VLF/LF lightning waveforms of CGs and ICs. A classification accuracy of 97% is achieved, which shows an excellent adaptability and automation(Zhu et al., 2021). The CNN models with different structures are proposed to perform the classification of VLF/LF signals generated by multiple lightning processes, including RS, PB, and NBE, etc. (Peng et al., 2019; J. Wang et al., 2020). It indicates that CNN has the potential to realize signal classification produced by various complex lightning discharge processes.

Although extensive efforts have been paid to improve the classification accuracy of lightning VLF/LF waveforms, towards to the development of high-performance short baseline VLF/LF lighting detection system, the following limitations still exist:

- Using the multi-parameter method, the classification accuracy of RS and IC in the LLS system has reached more than 90%. The classification accuracy may be further improved by optimizing thresholds of the multi-parameter method based on long term operation experience. However, since the VLF/LF waveforms produced by lightning leader discharges has more pulses and other high frequency components, it is difficult to determine thresholds involved in the multi-parameter method which can effectively discriminate different lightning events correlated to lightning leader propagation. Recently, it was found that the VLF/LF signals generated by NBE are wrongly identified as RS by the multi-parameter method(Leal et al., 2019; Lyu et al., 2015).
- The machine learning methods show promising performance in multi-object classification tasks, the challenges of applying machine learning methods in lightning VLF/LF waveform classification come from two aspects. Firstly, note that the data-driven nature of the machine learning methods means that the

performance is highly dependent on the balance and quality of the original dataset. An CNN model derived from imbalance data set is not reliable, because the model will tends to classify the objective waveform into the category which has the most samples in training dataset (Kaur et al., 2019). Secondly, since the characteristics of lightning VLF/LF signals can change in different regions, the accuracy of machine learning methods largely depends on whether the training dataset covers all possible variations of the objective waveform characteristics. Meanwhile, we need to note that most of the classification process by using machine learning methods acts like black box models, which makes it is difficult to ensure the classification accuracy of different lightning events. As discussed by Zhu et al., misclassification of RS signal can still occur by using SVM, although the characteristics of the misclassified waveform can be easily recognized manually. Since it is difficult to obtain the lightning waveforms in all regions of the world to expand the database, it is necessary to develop interpretable machine learning models to open the black box, which can reveal the classification process (Lipton, 2018) and assess whether the model is able to capture the essential characteristics of different types of lightning VLF/LF signals.

In this paper, a new interpretable CNN model which utilizes the class activation map (CAM) to represent the contribution of different waveform parts during the classification process is proposed. A four-class dataset including RS, PB, NB and IC is established for model training. The dataset is based on 17,441 waveforms recorded from five provinces in China with the latitude ranging from 29.1° to 33.5° and the longitude from 91.1° to 120.2°. The classification accuracy of the trained CNN is compared with that of the SVM model. The classification process of the four types of lightning waveforms is visualized by the CAM, which throws light on the relationship between the high-weighted waveform features and the physical process of leader discharge in lightning. The classification results are analyzed for the range of variation of the characteristic parameters of different waveforms in turn. The generalization of the proposed CNN model is test on another open dataset in Argentina used by Zhu et al. This paper is organized as follows: Section 2 introduces the data sources and the improved CNN network structure used in this paper. Section 3 shows the classification performance of the trained model and discusses the interpretability of the classification results. Section 4 discusses the universality of the CNN model, and Section 5 makes the conclusion.

2 Data and Methodology

2.1 Dataset

The dataset used in this paper comes from 17,441 lightning radiation waveform data recorded during 2019-2020. The measurement device is the VLF/LF electric field change meter (EFCM). The EFCM consists of an antenna and a digital data acquisition unit. The frequency band of the EFCM is 10Hz to 500kHz, with a sampling rate of 5Ms/s and a GPS synchronization error of less than 50ns (Y. Wang et al., 2020). The EFCMs were installed in five different provinces of China, including Hubei, Jiangsu, Zhejiang, Anhui and Tibet, as shown in Table 1. When the dataset covers waveforms in a variety of meteorological and terrain conditions, the model can be more generalized and the classification accuracy may be improved. The installation sites of these EFCMs have altitude between 190m to 4000m above sea level, within the longitude from 91.1° to 120.2° and the latitude from 29.1° to 33.5°. In order to improve the quality of recorded waveforms and exclude the impact of measurement noises, a combination of empirical mode decomposition (EMD) method and wavelet denoise method were used to pre-process the lightning radiation signal.

Table 1 Location of the deployed VLF/LF lightning waveform measurement meters

Location	Longitude	Latitude
Wuhan, Hubei	114.409	30.514
Sihong, Jiangsu	118.219	33.482
Wuxi, Jiangsu	120.256	31.618
Lishui, Zhejiang	119.656	27.976
Taizhou, Zhejiang	121.38	29.125
Hefei, Anhui	117.202	31.761
Anqing, Anhui	116.123	30.231
Lasa, Tibet	91.14	29.666
Linzhi, Tibet	94.373	29.636
Changdu, Tibet	97.179	31.146

Compared to the multi-parameter classification method, the machine learning algorithms can utilize the entire data information of time-resolved waveforms instead of several characteristic parameters. The SVM method was employed to classify the full VLF/LF waveform of lightning signals (Zhu et al. 2021). Considering the computational resources required for the deployment, the waveform is down sampled and divided into equal lengths, where each waveform slice is 100 μs in duration and contains 101 sample points. With the development of microprocessor technology in recent years, the main frequency of LS1043A ARM board we use has reached 1.6GHz, which significantly improves the ability to process waveform data per unit of time. In this paper, each waveform slice contains 2500 points corresponding to a time duration of 500 μs, which is beneficial to preserve the essential features of the original waveform. In the following discussion, each waveform slice is called as a sample. The dataset is constructed based on the prior knowledge of RS, PB, NB and active stage of IC, which can be found in reported literatures. We manually selected samples with the highest signal-to-noise ratio (SNR) and divided them into these four categories with a total of 8000 samples. To ensure a balanced dataset, each of the four categories contains 2000 samples. Note that our dataset is not classified by the polarity of lightning event, because the polarity can be easily intensified according to the polarity of the first pulse. It should be emphasized that this simplification will not affect the classification results discussed in the following parts. A 5-fold cross validation approach was adopted for the dataset, with the training set containing 6000 samples and the rest 2000 samples participate in the test.

2.2 Method

The CNN model performs feature extraction by convolving a convolution kernel with the input data. The convolution kernel is a weight matrix representing the features learned by the CNN model. The convolution kernel is usually initialized with random values, and the CNN model compares the output results with the true labels and updates the convolution kernel by backpropagation during the iterations. Thus, the convolution kernel can better match the core features of the data and improve the model's performance. When making VLF/LF lightning waveform classification, traditional CNN networks (plain CNN) have the following limitations:

- The same size of the convolutional kernel in each convolutional layer in plain CNN makes it difficult to handle the possible multi-scale features in waveforms.
- We use a high computing capability development board to process the data, the sample time in the dataset is longer and the sample includes more information, requiring more convolution layers to fully extract these features. However, the backpropagation process calculates derivatives in chains to update network parameters. The gradient information may vanish gradually when simply increasing the number of convolution layers. It can also cause model degradation and make the model difficult to converge (He et al., 2016).
- The CNN model flattens the features extracted from the convolutional layers into a one-dimensional vector, and output the classification results by means of full connection. It is difficult to obtain which part of the waveform determines the classification results of the model, which makes it impossible to judge whether the

classification process of the model is reliable.

In order to fix those issues, we proposed an interpretable CNN model with improved performance in feature extraction and convergence speed. The model includes a CNN classifier and a visualization module as shown in Figure 1.

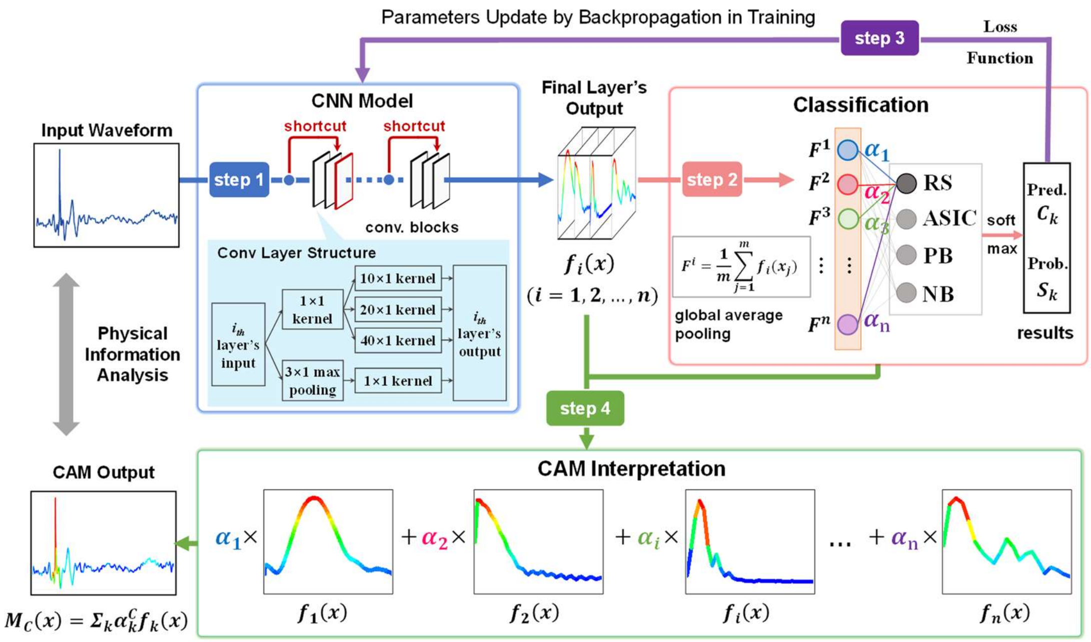


Figure 1 Structure of the proposed interpretable CNN model

a) ***The CNN classifier:*** The proposed CNN classifier takes waveforms as input and gives out classification results with probabilities. Step 1-3 describes how the CNN classifier works and self-upgrades iteratively in training.

In step1 the waveform is fed into the CNN model and the high-dimensional feature maps are obtained. Compared with the plain CNN, the proposed CNN model adopts shortcut connections and parallel convolution kernels. The CNN model contains two convolution blocks, which is formed by stacking three convolutional layers. In each block, part of the input data is directly transferred to the second layer of the block through a shortcut connection. The shortcut connection aims to solve the problem of model degrading in multi-layer networks and accelerates the convergence in training. In each convolution layer, the convolutional kernels with the size of 40, 20, 10 and 1 are introduced in a parallel structure. The kernels with the size of 40, 20 and 10 give the model a more various feature matching range after multiple layers, enabling the extraction of long-scale waveform features. The kernels with the size of 1 ensure that the model can also capture detailed waveform features. Each layer can be expressed as:

$$f^l = b^l + \sum_{i=1}^{N_{l-1}} conv1D(w^l, f_i^{l-1}) \qquad (1)$$

Where x^l is defined as the input of layer l, b^l is defined as the bias layer l , f_i^{l-1}

is the i_{th} output part of layer *l-1*, w^l is the multi-size convolution kernels at layer *l*, N_{l-1} is the number of output in layer *l-1*.

In step2, the feature vector is obtained through the global average pooling based on the feature maps produced in step 1. Compared with the plain CNN methods, which flatten the high-dimensional feature maps as feature vectors, the proposed model uses the global average pooling to form the feature vectors and greatly reduces the computations of the model.

The classification probability of the waveform is computed by the fully connected layer and the SoftMax function. The probability S_c that a waveform belongs to a category *c* can be obtained from equation (3):

$$S_c = \Sigma_i \alpha_i^c F^i \qquad (2)$$

Where α_i^c represents the contribution of feature map $f_i(x)$ to model's classification result of category *c*.

During the model training, the model's classification will be compared with the true label of the waveform by the loss function as shown in step 3. The result is referred as the loss value in training. The model uses the back propagation algorithm to make the loss information flow backward to update model parameters, which can be expressed as:

$$\frac{\partial u^n}{\partial u^j} = \sum_{i:j\in Pa(u^i)} \frac{\partial u^n}{\partial u^i}\frac{\partial u^i}{\partial u^j} \qquad (3)$$

Equation 5 describes how to calculate the gradient of an output node u^n (such as the loss value) over several input nodes from u^1 to u^j to achieve gradient descent update of the parameters. Where u^i refers to the intermediate nodes in all possible paths (Pa) from u^n to u^j. The gradient is essential for the gradient descent optimizing method during parameters update.

b) ***Model Interpretation:*** In order to open the black-box of the CNN model, we introduce the CAM method in the proposed CNN model as shown in step 4. The CAM method multiply the weight vector produced in step 3with the high-dimension feature maps produced in step 1 to obtain a class activation map (CAM) which can mark the important waveform features in the classification. The CAM values for the class can be defined as:

$$M_C(x) = \Sigma_i \alpha_i^C f_i(x) \qquad (4)$$

We denote $M_C(x)$ as the CAM value of the waveform under category *c*. By using heat maps the CAM value provides a direct indication of the importance of each datapoint *x* to the classification result of category *c*.

Due to the model structure difference, CAM method is not appliable on traditional machine learning methods like the SVM. We use the SHAP method to visualize the important waveform features in the classification of traditional machine learning methods for comparison. SHAP method derives from cooperative game theory, which provides global and local interpretability of the features(Lundberg & Lee, 2017). The SHAP value is based on the marginal contribution of the features amongst all the feature arrangements. In waveform classification, we regard each waveform datapoint as a feature, the SHAP value can be expressed as:

$$\varphi_j(val) = \Sigma_{S \subseteq x_1,\cdots,x_M/x_j} \frac{|S|!\,(M-|S|-1)!}{M!}(val(S \cup x_j) - val(S)) \quad (5)$$

Where the x_j is the j_{th} feature and refers to the j_{th} point of input waveform. φ_j is the contribution value of x_j to the classification result. S is the subset of features and M is the total number of features. The value function $val(*)$ refers to the model's classification results in machine learning. In the following section, SHAP is referred to the φ_j and used to describe the contribution of different waveform features to the classification results in traditional machine learning methods.

2.3 Model Training

In this paper, the proposed CNN model is deployed on a Tesla A100 graphics card and the training framework is Pytorch 1.9.1. The hyperparameters required for training are shown in Table 2.

Table 2 Hyperparameters used for training

Hyperparameter	Value
Batch Size	48
Epoch	40
Loss Function	CrossEntrophy
Optimizer	Adam
Learning Rate	0.01
Momentum	0.5

Batch Size means that the model is fed with 48 data samples at a time in model training, and the parameter Epoch specifies that the model performs a total of 40 forward calculations and back propagation processes. The loss function used in the model is the CrossEntrophy function and the Adam is used as the optimization algorithm. The Learning Rate and Momentum together control the convergence rate and the efficiency of the model, which are set to 0.01 and 0.5 respectively after pre-test. The model was trained under the above conditions, and the training loss and classification accuracy are shown in Figure 2. Due to the introduction of the shortcut connection to solve the vanishing gradient problem, the model converges faster and reaches convergence after only five epochs, with an overall classification accuracy of about 98.5%.

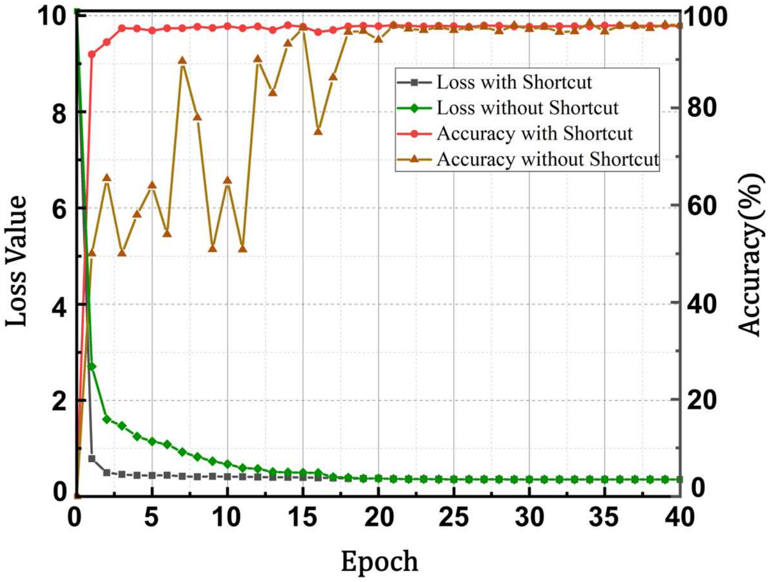


Figure2 Loss and accuracy changes of plain CNN model without shortcut and improved CNN model with shortcut in training

3 Results

3.1 Comparison for classification results

After training, we compared the classification results of the proposed CNN method with other machine learning methods such as SVM and RF under the same dataset. The feature vector used for training SVM is obtained by data down sampling method (Zhu et al., 2021) and the amplitude-frequency features were extracted as feature vectors when training the RF model (Nassralla et al., 2017) . The performance of these methods is shown in table 3:

Table 3 Comparison of the results of different models

Method		CNN				SVM	RF
Metrics		Accuracy	Precision	Recall	F1	Accuracy	Accuracy
Class	RS	96.8%	1.00	1.00	1.00	90%	88.3%
	PB	100%	0.94	0.97	0.96	91%	83%
	IC	100%	1.00	1.00	1.00	86.20%	84.5%
	NB	97.8%	0.97	0.94	0.96	92%	92.7%

Table 3 shows the comparison of classification accuracy in the four kinds of waveforms. For waveforms with short duration such as RS and NB, both CNN method and traditional machine learning methods achieve good classification results in classification accuracy, while the CNN model has an improvement of about 6%. However, for waveforms like PB and IC which last longer and is more difficult to classify, the CNN method shows a significant improvement of up to 17% in accuracy compared to traditional machine learning methods. The performance of SVM and RF is not as expected as that in the original literature. This may be due to the fact that the dataset we used contains longer slices of waveform which can preserve more waveform features. Therefore, the hyperparameters used in the original literature may be no longer suitable and need to be adjusted. In contrast, the CNN model supports raw waveforms as inputs, without the need of manual adjustment of hyperparameters after changing datasets, which can bring stronger robustness. In addition, the CNN model performs well on metrics like precision, recall and F1, proving that the model does not have an imbalance problem.

3.2 Model interpretability analysis based on CAM visualization

The proposed CNN model achieves higher performance in all kinds of waveforms, which may be related to a better understanding of the physical process. To investigate the feature the model has learned, we use CAM visualization method mentioned in section 2.2 to estimate which part of the waveform lead to its classification result. For comparison, the SHAP method is used to mark the waveform which owns higher contribution to affect the classification results in SVM(Ribeiro et al., 2016). In this section we will discuss the classification process of the four typical VLF/LF waveforms in both cloud to ground flashes (CG) and intracloud flashes (IC), including return stokes (RS), active stage of IC (ASIC), preliminary breakdown (PB) and narrow bipolar pulses (NB). The accurate classification of RS helps to distinguish between CG and IC events and can improve the location efficiency of LLS. The classification of PB, NB and ASIC is important for further research of the initiation mechanisms of the lightning discharges.

3.2.1 Return Stroke Produced by CG

The LLS allows for real-time detection of return strokes (RS) in cloud to ground flashes (CG) due to the strong and widely spread VLF/LF waves of RS, which is generated by the propagation of high amplitude currents in the lightning channel. The waveshape and amplitude of RS are closely related to factors like current strength, propagation speed and propagation path etc. Compared to other stages of the lightning discharge, the velocity of RS is high and the current of RS is strong. Therefore, the LLS often use the amplitude or pulse width of the electrical field waveform to identify RS. The traditional multi-parameter classification method concludes that the pulse width of the RS waveform is typically between 10 and 200 us. However, recent reports indicate that using pulse width as a criterion can easily misclassify several kinds of intracloud lightning pulses as RS (Biagi et al., 2007; Leal et al., 2019; Nag et al., 2014). In recent years, scholars have already tried using machine learning methods like SVM to perform RS/IC classification (Zhu et al., 2021). This section applied the CNN model trained in Section 3.1 for RS classification. By analyzing the classification process of our CNN model under several typical RS cases, we explored the key physical features the model learned and compared it with the SVM method.

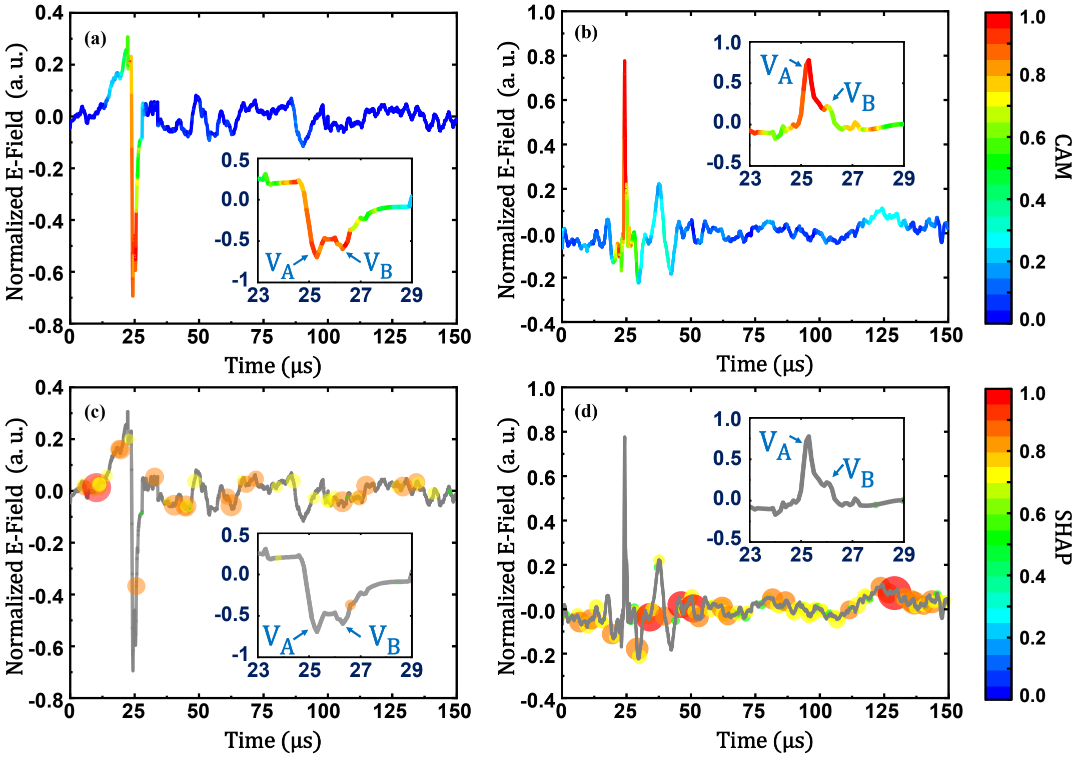


Figure 2 Visualized classification results of negative RS waveform #190912135803-001RS and positive RS waveform #190912135803-003RS waveform. **a)** Visualized CNN classification result based on CAM for the negative RS case with a detailed demonstration for the main pulse part **b)** Visualized CNN classification result based on CAM for the positive RS case with a detailed demonstration for the main pulse part **c)** Visualized SVM classification result based on SHAP for the same case and detailed demonstration as (a) **d)** Visualized SVM classification result based on SHAP for the same case and detailed demonstration as (b)

Figure 3(a) shows the classification result of the CNN model for a negative RS, which was recorded at 13:58:03 September 2019 at Anqing, Anhui, China. We define data points with CAM weight values above 0.5 as hotspots during classification. The pulse width of this case is about 11 μs, which is at the lower threshold according to the multi-parameter method, leading to great possibility for misclassification. However, the CNN model gives out a hotspot region between 25μs and 27μs, which means the CNN model accomplished the classification mainly by the main peak part with a duration of only 2μs. It can be seen from this part that the main pulse contains a sequential double-peak characteristic with a primary peak VA and a subpeak VB. Based on observation results, Le Vine et al. conclude that the subpeak structure of the RS waveform is related to the geometry change of the lightning channel (Le Vine, 1980). Cooray et al. propose that the abrupt changes in current amplitude or channel development velocity will result in the subpeak structure in the VLF/LF waveform (Cooray & Lundquist, 1985). Figure 3(a) demonstrates that the CNN model successfully captured the double-peak characteristic of the RS waveform, which represents a key part of the physical process of RS.

Figure 3(c) gives the SVM classification result for the same negative RS waveform. The orange and red circles represent the high weight points given by the SVM with a SHAP value greater than 0.5. The high weight points distribution shows that the SVM model is able to depict the profile of the waveform. It can be inferred from this that the SVM method may classify the RS waveform by marking the high-amplitude part of the RS waveform. Although the SVM model was also able to correctly classify this RS event, it failed to capture other physical information of the RS waveform.

The classification results for another positive RS are given in Figure 3(b) and Figure 3(d). The hotspot region in Figure 3(b) shows that the CNN model also captured the two-peak feature of both PA and PB, which means a better understanding about the correlation between the RS waveform and physical process such as the change in the channel development velocity and the channel geometry. However, the SVM method fails to classify this positive RS waveform. As can be seen from Figure 3(d), the high weight points given out by the SVM model also tends to depict the entire waveshape. But due to the bipolar pulse following the main pulse of this waveform, the SVM method failed to mark the main pulse, which resulted in a 42% probability for RS while an 82% probability for NB. The CNN model proposed in this paper not only marks the true main pulse part of the RS waveform but also captured the double-peak feature in the main pulse. Compared to multi-parameter method, the proposed CNN model can overcome the problem that an applicable general criterion for parameters like pulse duration is difficult to be determined. As for traditional machine learning method like SVM, the anti-disturbance capability of the proposed CNN model also gets improved, which means stronger robustness.

3.2.2 Active Stage of ICs

Intra-cloud lightning discharge (IC) occurs in a single storm cloud or between different storm clouds. The intra-cloud lightning discharge can be divided into wo stages, including the active stage and the final stage (Bils et al., 1988). The VLF/LF radiation signal generated during the active stage of intra-cloud lightning (ASIC) presents a sequence of pulse activities. A variety of transient processes appears in the following final stage, including the narrow bipolar events, the stepped leader, the J process and the K process, etc. (Rakov & Uman, 2003). The VLF/LF radiation signal of the final stage of ICs is usually not used to identify the intra-cloud lightning events, because it owes an overlapping amplitude range with the RS. Conversely, during the ASIC, repetitive VLF/LF electric field pulses can be detected. These pulses are characterized by low amplitude and unipolarity and are related to the stepped growth of the negative leader, which are applicable to distinguish the IC and CG(Brunner, 2016). However, the statistics of characteristics of the pulses during the ASICs are rarely reported. In this section, attempts were made to demonstrate that an interpretable CNN model can be used as an effective approach for the classification of ICs events.

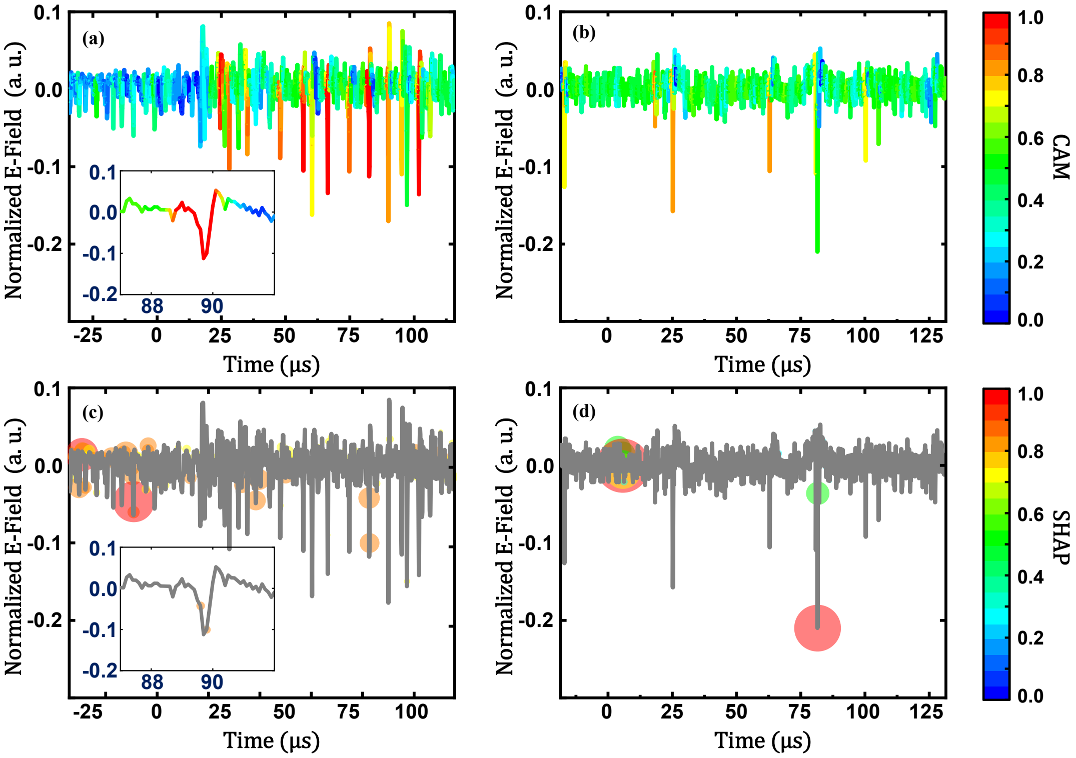


Figure 3 Visualized classification results of negative IC waveform #190818114117-001IC and #190817184223-001IC. **a)** Visualized CNN classification result based on CAM for the first case with a detailed demonstration for the one single pulse **b)** Visualized CNN classification result based on CAM for the second case with a detailed demonstration for one single pulse **c)** Visualized SVM classification result based on SHAP for the first case **d)** Visualized SVM classification result based on SHAP for the second case

Figure 4(a) shows the visualized CNN classification result for a pulse train during the active period of the IC lightning. This waveform is captured at 11:41:17 18th August 2019 at Lishui, Zhejiang, China. The electrical waveform at this stage consists of a sequence of pulses. The tail of each single pulse is followed by a small, slowly changing polarity-reversed process (Krider et al., 1975). The pulses repeat slowly at this stage, with an average pulse interval of 10.7μs in this case, which is consistent with the existing observations (Gomes & Cooray, 2001; Krider et al., 1975). In Figure 4(a), the hotspot region given out by the CNN model mainly contains two parts. Firstly, the model focuses on the pulse peak and its subsequent polarity-reversed part. The mean CAM value of this region is greater than 0.8. It shows that the CNN model provides a good understanding of the causal relationship between the main peak and the subsequent polarity-reversed waveform and treat them as a whole part. Secondly, it should be noted that the waveform between the two pulses is mainly background waveform, but the model still gives out CAM values of 0.5 to 0.6, significantly higher than the CAM values of the background waveform outside the active stage. It can be inferred that the model not only captures the characteristics of single pulses of ICs, but also find the pattern of continuously pulse repeat during the ASIC. In Figure 4(c), the SVM method is used to classify the same case. The circles in Figure 4(c) represent the high weight points given out by the SVM with a SHAP value greater than 0.5. The high weighted points mainly locate near the peak of the pulse. This suggests that although the SVM method can also capture the peak of the pulse, it fails to capture the causal relationship of main peak and the succeeding part.

Figure 4(b) and Figure 4(d) compare the classification results of CNN and SVM method for another IC waveform. For the CNN model, the hotspot region is similar to that in Figure 4(a) which also concentrate on the single pulses and the pulse intervals. However, the SVM method misclassifies the waveform as RS. Figure 4(d) shows that the high weighted points of the SVM model mainly locates around 0μs and 75μs. The waveform around 0μs is mainly the background electric field and the waveform around 75μs has the highest pulse peak. The results show that the SVM model mainly focus on the peak of the electric field pulses. Besides, since the SVM model mainly focuses on the pulse peaks, it makes the high weight points in this case locate around the largest pulse, leading to the misclassification as RS. By the comparison, it can be concluded that the CNN model is able to learn the detailed features like the temporal relationship between the first peak and subsequent polarity-reversed part of pulses, and is also able to effectively identify the macro features like the repetition of pulses in active stage of ICs.

3.2.3 Preliminary Breakdown Pulses

The preliminary breakdown (PB) is the initiation and development of the leaders in cloud, which is considered to be the initial stage of the lightning. The VLF/LF electric field waveform generated by PB is composed of consecutive bipolar pulses with a total duration of microseconds. It is concluded from theoretical simulations that the waveform of the PB process has a similar physical mechanism to that of the NB, which is probably related to the consecutive stepped elongation of the negative leader channel within the thundercloud(Da Silva & Pasko, 2015). The multi-parameter method usually utilizes the SNR to make classification. It is considered to be a PB process while at least three consecutive bipolar pulses are found with peaks twice the average noise level or more (Nag & Rakov, 2008). However, according to the discussion in section 3.2.2, the waveforms during the ASICs are also characterized by repetitive bipolar pulses, which makes it difficult to achieve an accurate distinction between the PBs and ASICs using multi-parameter method. In this section, the visualized result of the CNN model for PB classification is analyzed to illustrate the CNN model's ability to capture the physical features of PB waveforms, which improves the model's accuracy and robustness.

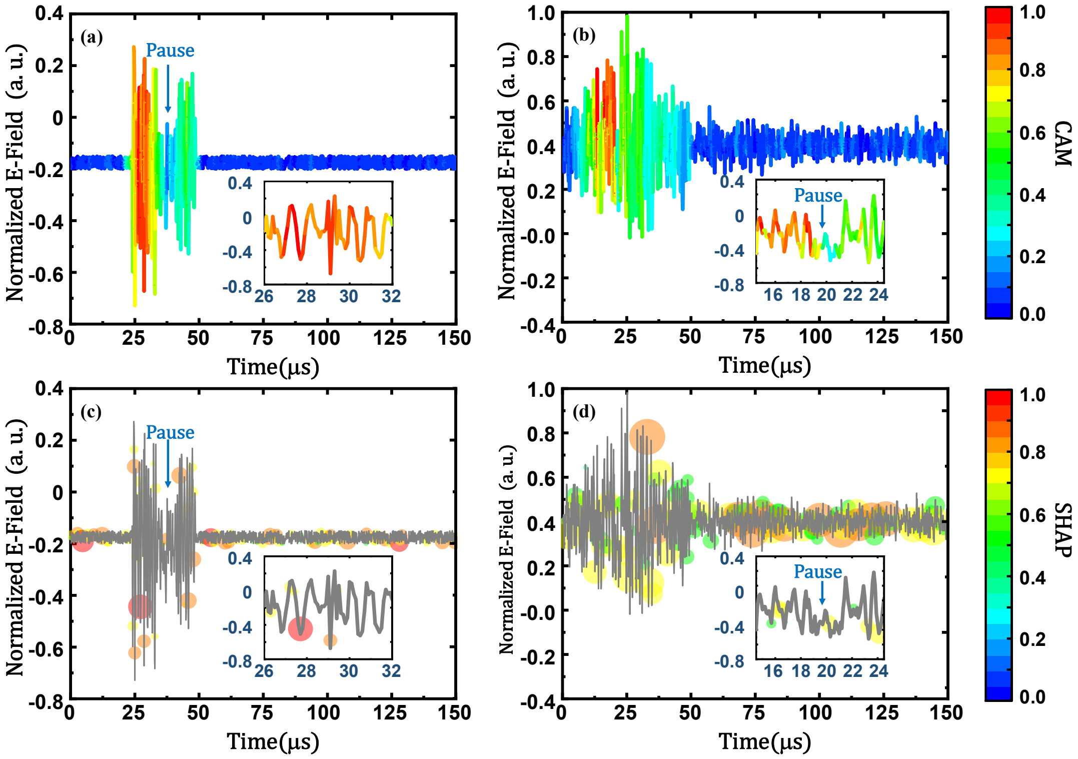


Figure 4 Visualized classification results of PB waveform #190727133157-001PB and #1907271314423-001PB. **a)** Visualized CNN classification result based on CAM for the first PB case with a detailed demonstration for the main pulse and the pause part **b)** Visualized CNN classification result based on CAM for the second PB case with a detailed demonstration for the pause part **c)** Visualized SVM classification result based on SHAP for the first case **d)** Visualized SVM classification result based on SHAP for the second case

Figure 5(a) shows the visualized CNN classification result for an PB waveform, which is captured at 13:31:57 27th July 2019 at Wuxi, Jiangsu, China. The hotspot region given out by the CNN model covers the entire period between 23 and 54μs in which the pulses exist. However, it is obvious that two partitions with different CAM values exist during this period. The first partition is between 26 and 32μs where the CAM values are all above 0.8. It can be seen from Figure 5(a) that the waveform within this partition is notably characterized by significant continuous bipolar oscillations. The second partition is between 38 and 54μs where the CAM values range between 0.4 and 0.6 and also has significant continuous bipolar oscillations. It is interesting that in the time from 32 to 38μs between the two partitions, there is a pause interval where the CAM values are less than 0.3. Figure 5(a) shows that the amplitude is low and bipolar oscillation is not obvious during the interval. The above analysis shows that the model in this paper is able to adaptively mark intervals that match the PB characteristics based on the bipolar oscillation frequency and amplitude characteristics of the waveform. Figure 5(c) shows the visualized SVM classification result for the same case. It can be seen from Figure 5(c) that the high weight points distribute on both the background waveform and the pulse part, which indicates that the SVM completes the PB identification by depicting the overall profile of the waveform without an understanding of PB’s core physical features.

To further compare the proposed CNN model with the SVM model, another case of PB is shown in Figure 5(b) and Figure 5(d). In Figure 5(b) the CNN model gives a similar distribution of hotspot regions as in Figure 5(a). The CNN model accurately marks the pulse part which is also divided into two partitions by a pause interval, and the duration of the pause interval in Figure 5(b) is shorter by about 1μs compared to figure 5(a). This represents a better ability of the CNN model to adaptively classify continuous pulses that conform to bipolarity, with better robustness. The above phenomenon means that the CNN model is able to adaptively find PB-like waveforms, and even a very short non-PB interval will cause the CAM value to drop sharply. In comparison, Figure 5(d) shows that the SVM method misclassifies this case, with the highest weight points distributing at the start of the waveform around 30μs. Besides, the other high weight points mainly exist in the negative part of the waveform, leading to the misclassification. It can be inferred that the CNN model has higher classification accuracy compared to traditional machine learning methods for the ability to recognize temporal features like the continuous bipolar pulses.

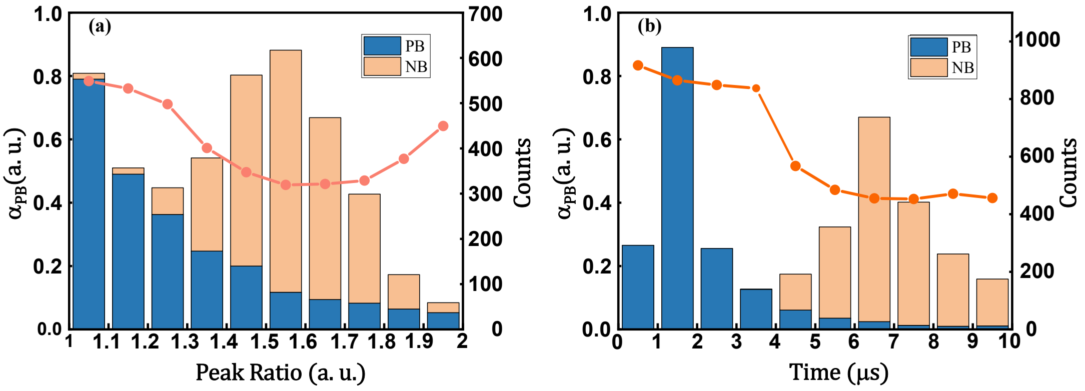


Figure 6 (a) relationship of CAM values α_{PB} and peak ratio P (b) relationship of CAM values α_{PB} and pulse duration T

To demonstrate our model's ability to help to find adequate threshold for multi-parameter classification. The average CAM values α_{PB}, peak ratio P and pulse duration T are estimated and compared with the results of similar waveforms like NB. We define a single pulse as a segment of the waveform between two zero crossing points, which must contain a polarity change. The α_{PB} refers to the average CAM value given by the model when the waveform is considered to be a PB. The peak ratio P is the absolute value of the first and second peak amplitude ratio, and the pulse duration T is the time interval between two crossing points. The α_{PB} -P relationship is given in Figure 6(a). The result shows that the bipolar peak ratio of the PBs waveform ranges between 1 and 2, which is in agreement with the range of peak ratios of the NB. The difference is that 63% of the PBs have a P of less than 1.3, while 81% of the NBEs have a P between 1.3 and 1.7. It is notable that when the P is less than 1.3, the α_{PB} is greater than 0.5 and decreases as the P increases. This suggests that the more P is close to 1, the more likely the pulse considered to be PB in our model, which is consistent with the simulation results of Silva et al. (Da Silva & Pasko, 2015). However, when the P is greater than 1.8, the α_{PB} increases as the P increases. This suggests that the peak ratio cannot be used as an effective way to distinguish PB from NB. This may be because that the P changes as the conductivity of the leader channel changes in which the PBs radiation source locate, thus deviating from the theoretical result of Silva(Kašpar et al., 2017). Figure 6(b) demonstrates the relation of α_{PB}-T, where T is within 10 μs for both PB and NB. 91% of PB had T of less than 4.0 μs, while all of NB have T between 4.0 and 10.0 μs. As can be seen from the trend of α_{PB}, the proposed model suggests that the shorter the pulse duration is the more likely the pulse is to be a PB, especially when the pulse duration is less than 4μs. It should be pointed out that as the T increases, the α_{PB} gradually decreases to around 0.4. According to the results, overlaps exist in the parameter distribution of PB and NB, which lead to the difficulties to set an adequate threshold for multi-parameter classification. However, the turning points of α_{PB}-P and α_{PB}-T is generally consistent with the actual peak ratio and pulse duration distribution of PB and NB. This indicates the CAM values from the proposed model is helpful in threshold determination.

3.2.4 Narrow Bipolar events (NBs)

Narrow bipolar event (NB) is a special type of intracloud discharge, often occurring in isolation from other discharge events. The amplitude of NB is usually high, which can be close to that of RS(Rakov & Uman, 2003; Smith et al., 1999). The pulse duration of NB is short, ranging from 2 to 20μs (Jacobson and Light 2012; Wu et al. 2014). Increasing evidences indicate that NB may be the initiating process for other lightning discharge events. It is demonstrated through simulation that the electric field waveform of the NBE is related to the abrupted elongation of the initial negative leader channel in the thundercloud(Da Silva & Pasko, 2015). Because of the distribution overlap of several essential characteristic like the amplitude and pulse width between NB and RS, the NB is thus an important factor affecting the accuracy of RS classification in LLS.(Leal et al., 2019; Nag et al., 2014). In this section, the visualization of which part of the NB waveform causes the model to make the correct classification will be analyzed.

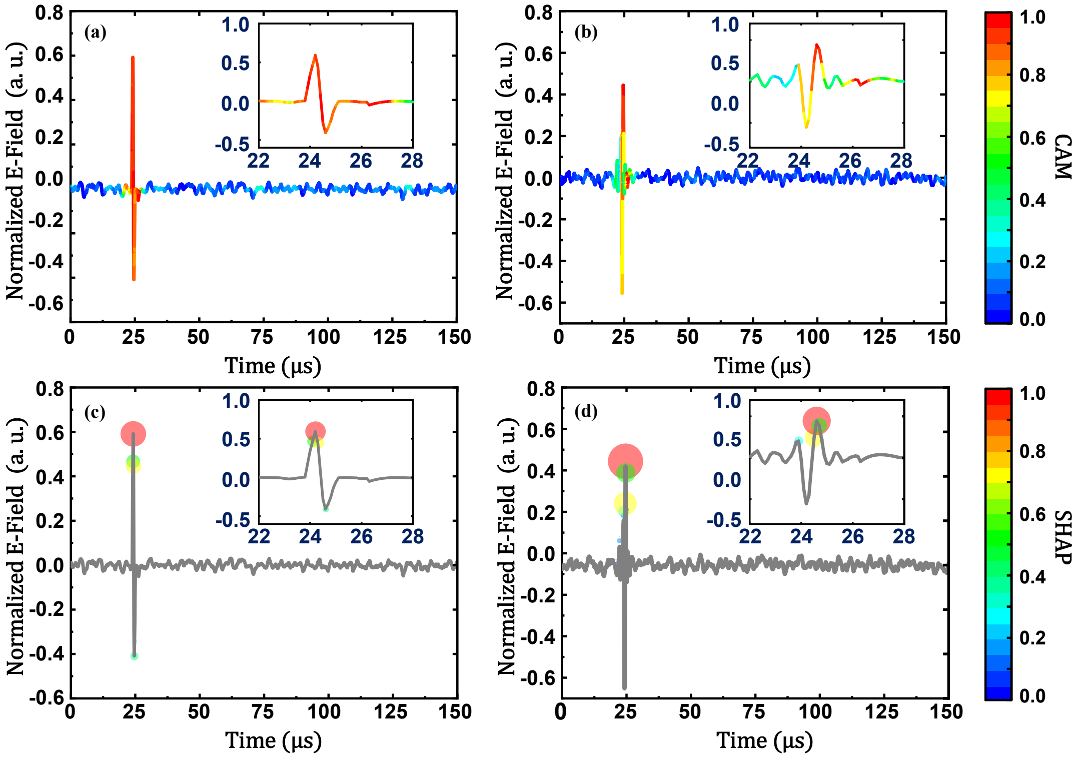


Figure 7 Visualized classification results of NB waveform #190818114306-001NB and #190818114919-001NB. **a)** Visualized CNN classification result based on CAM for the first NB case with a detailed demonstration for the main pulse part **b)** Visualized CNN classification result based on CAM for the second NB case with a detailed demonstration for the main pulse part **c)** Visualized SVM classification result based on SHAP for the first case and detailed demonstration as (a) **d)** Visualized SVM classification result based on SHAP for the second case and detailed demonstration as (b)

Figure 7(a) shows the visualized CNN classification result for an NB waveform, which is captured at 11:43:06 18th August 2019 at Anqing, Anhui, China. It can be seen that the proposed CNN locates the main pulse part of this waveform within 24.5±1.5μs accurately, which contains both positive and negative peaks as well as a steep polarity change process. It is notable that the average CAM value of the main pulse is higher than 0.8, while the average CAM value of the other part is less than 0.2, indicating that the CNN model focuses on the waveform pulse and pays less attention to the background waveform. The CAM value difference shows that the CNN model captures the feature of isolation of the NB waveform, with attention focused on the pulse part in the waveform which contains the most information of the waveform. Figure 7(b) shows the visualized SVM classification result for the same case. Compared to the CNN model, the SVM model also focuses on the pulse part and the high weight point is distributed at the positive and negative peak tops, with almost no high weight point distributed in the background part. Figure 7(c) shows that the weight values positive is positively correlated to the pulse magnitude, indicating that the SVM model may classify waveforms by identifying the high amplitude parts in the waveform. It is important to point out that the two core characteristics of the NB require the model to be able to recognize time-dependent feature. The short pulse duration feature requires the model to be able to determine if the pulse presents for a certain period of time. The bipolar pulse feature requires the model to be able to determine if positive and negative peaks appear in succession. As depicted in 2.3.2, the parallel design of multiple convolutional kernels in our CNN model allows the model to capture features at different time scales during the training process. Due to the variable scale of the features extracted by the CNN model, the model can learn the temporal causality features in waveforms during the learning process. Therefore, the hotspot region given out includes the entire main pulse part. In contrast, the SVM model only focuses on the high amplitude point distribution of the waveform and therefore may lead to misclassification of some NB waveforms. Figure 7(b) and Figure 7(d) show another case of NB. The bipolar pulse in this case is located around 23.9 to 24.9μs, but there is a large positive disturbance at 23.5μs with a peak ratio of approximately 0.5 to the main pulse. Figure 7(b) shows that the CNN model successfully classifies the waveform, with the given hotspot region distributed between 23.8-24.9μs, which is coincides with the main pulse duration. It should be noted that the CAM value of the positive part disturbance is less than 0.5. However, the SVM model is unable to classify this case correctly. As can be seen from Figure 7(d), the SVM model only locates the positive peak top of the main pulse as well as the positive peak top of the disturbance, ignoring the negative period of the main pulse, and thus makes incorrect judgments as a result. As can be seen from the above comparison, the SVM model only classifies waveforms by the distribution of high amplitude points, which lacks consideration of temporal correlation. The classification of SVM is not supported by physical process and is prone to misclassification. The CNN model takes into account the temporal correlation patterns in the timeseries data at an adaptive scale during the classification process, and is more likely to capture the core features of the NB waveform, providing higher classification accuracy and robustness.

To further demonstrate that the CNN model can help to find threshold to distinguish similar waveforms like RS and NB, we calculated the peak ratio P, pulse duration T and average CAM values α_{NB} of the NB and RS waveforms in the dataset. The definitions of P and T are identical to those in 3.2.3. The α_{NB} refers to the CAM value given by the model when the waveform is considered to be a NB. The result is shown in Figure 8.

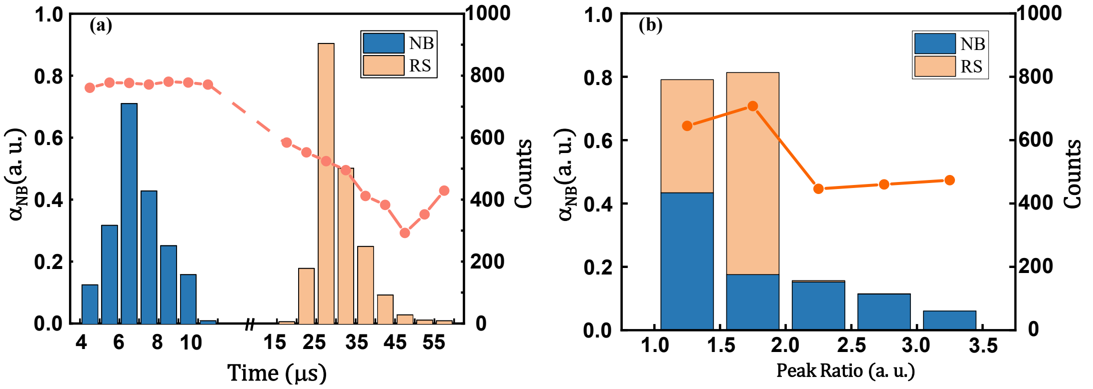


Figure 8 (a) relationship of CAM values α_{NB} and pulse duration T (b) relationship of CAM values α_{NB} and peak ratio P

Figure 8(a) shows the relationship between the average CAM value α_{NB} and the pulse duration T. When the pulse duration T is between 2-5.5μs, the average CAM value α_{NB} remains high within 0.75-0.8. As the T increases, α_{NB} drops rapidly to a minimum value of 0.22. It can be concluded from figure 8(a) that the model suggests that the pulse duration of the NB should not exceed 15μs. Figure 8(b) shows the relationship between the average CAM value α_{NB} and the peak ratio P. The peak ratio quantifies the degree of bipolarity of the waveform. When the P is between 1 and 2, the bipolarity of the waveform is more obvious and the average CAM value is at a high level between 0.6 and 0.8. When the P is greater than 2 and increases further, the bipolar characteristics of the waveform can be considered to have gradually disappeared and the unipolar characteristics become prominent. The CAM value drops steeply and remains at a low level of around 0.4. Figure 8(a) and (b) show that the turning points of α_{NB}-P and α_{NB}

-T is generally consistent with the actual peak ratio and pulse duration distribution of RS and NB.

4 Discussion

Compared to plain CNN models, our model employs the shortcut connection to improve the rate of convergence, which enables the model to deal with longer input waveforms. The multi-size kernels of our model can capture multi-scale temporal features of VLF/LF waveforms. In section 3.2, we demonstrate that the improved model can extract physical information of VLF/LF waveforms which is related to different lightning discharge processes. It means that the classification accuracy of our model is much less dependent on the dataset, and it can be extended to recognize the lightning VLF/LF waveform recorded from other regions.

The open dataset measured at Córdoba in central Argentina (Zhu et al., 2021) is employed as a test dataset here to further exam the portability of our model. The waveforms in the original dataset are classified into four types, including +CG, - CG, + IC and – IC. We ignore the impact of polarity on the classification accuracy since it is convenient to be recognized. The original dataset is reclassified into two groups of RS and IC. We down sample these waveforms to the sample rate of 5MS/s, which is corresponding to our devices. Since the input waveform of our model must possess 2500 points, we manually add noise data to the waveforms meet the required length. Higher classification accuracy is obtained by using our model than the result reported by Zhu et al. The classification accuracy of RS in our model is 99.41%, while it is 97% by using SVM. The classification accuracy of IC in our model is 97.38%, while it is 97% by using SVM. Note that the classification accuracy of IC here equals to the sum of PB, NB, and ASIC for convenience of comparison.

It should be emphasized that traditional machine learning methods can only give out the probability for different categories. The proposed model can visualize the contribution of different parts of the waveform to the classification result. One can observe in figure 9 that our model can effectively capture the physical features of waveforms which cannot be classified correctly by the SVM. Fig. 9 (a) shows the CAM visualization of a RS event. This case is misclassified as an IC by the SVM, which may be caused by the unexpected bipolar oscillation around 30μs according to the discussion of Zhu et al. It can be seen from figure 9(a) that the proposed model marks the dual peaks of the waveform which is believed to be an instinct feature of RS events. Moreover, the unexpected bipolar oscillation is neglected by our model, since the corresponding CAM are less than 0.5 there. Figure 9(b) shows the CAM visualization of a NB event. It is misclassified as a CG by the SVM for the positive disturbance around 27 μs according to Zhu's report. It can be seen in figure 9(b) that the steep polarity change process is marked with high CAM values, which is the key feature in the NB waveforms. The positive disturbance is neglected with the CAM values less than 0.5. According to these comparisons, the proposed CNN model is more effective in extracting the key features related to the physical process. These key features are relatively invariant in different regions. Therefore, the CNN model is able to accurately classify data from different regions.

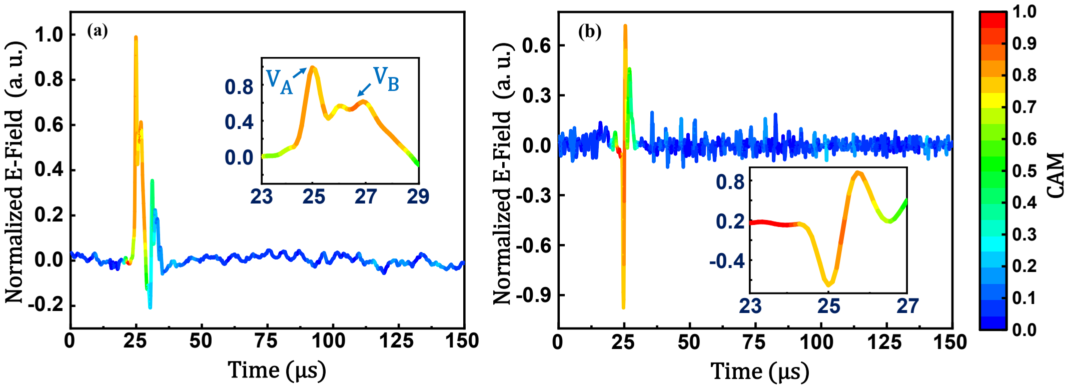


Figure 9 The CAM values of two misclassified waveforms by the SVM in the open dataset. (a) a RS waveform misclassified as IC. (b) an IC waveform misclassified as RS

5 Conclusion

In this paper, the main conclusions are summarized as follows:

（1）In this paper, an interpretable CNN model for VLF/LF lightning waveform classification is proposed. The proposed model uses multi-scale convolutional kernels to enhance the ability to capture local waveform features. The output of the final convolutional layer and the fully-connection weights are used to visualize the contribution of different waveform parts to the classification result. A shortcut connection is built in the proposed CNN model to promote the convergence speed and make the model capable of waveforms with higher sampling rate. Based on 8000 waveforms recorded in five provinces in China, the four-type classification of waveforms including RS, ASIC, PB and NB, is achieved with an accuracy of 98.5%, which is better than the traditional SVM and RF methods.

（2）Based on the distribution of the high-contribution waveform parts in the classification process, we analyzed the correlation between the model's focused features and the lightning discharge process. The model considers the double peak structure superimposed on the main pulse as the main feature of the RS, which is mainly caused by the abrupt change of the current or the branches of the lightning channel. The proposed model is able to identify the separated, repetitive pulses generated by the ASIC event, which are associated with the stepped growth of the negative leader in thunderstorms. The continuous bipolar pulse train is considered as the main feature of PB by the model, which is generated by the continuous development of the initial leader in thunderstorms. The model in this paper is also able to identify the narrow bipolar pulse generated by the sudden elongation of the initial leader in NB events. This indicates that compared to traditional machine learning methods, the model in this paper extracts features which are in line with the human experts in the VLF/VF waveform model classification process.

（3）We analyzed the relationship between the average waveform weights given by the model and the pulse duration, peak ratios. For NB and PB events with similar physical mechanisms, the weight value α_{PB} for PB events varies in a U-shape with the increase of the peak ratio. When the pulse duration T_{width} is greater than 4.0 μs, α_{PB} decreases monotonically with the increase of T_{width}. This indicates that the pulse duration is more suitable than the peak ratio to distinguish the NB and PB waveforms. Compared to the RS, the weight value α_{NB} for NB does not vary significantly (which is between 0.4 and 0.7) with the peak ratio. And when the pulse duration T_{width} is greater than 5.0 μs, α_{NB} decreases significantly with the increase of T_{width}, which indicates that the pulse duration can better solve the problem of easy confusion between RS and NB events in the lightning location system.

（4）We validated the model in this paper using an open source dataset reported in literature, which has a total of 32,754 samples from central Argentina. The model in this paper achieved an accuracy of 98.39%, which is better than the result using the SVM according to the literature. Based on the contribution weights obtained in this paper, it can be seen that the model in this paper considers the double-peaked structure superimposed on the main pulse as the key feature of the RS, which avoids the influence of unexpected waveform oscillation and NB waveforms on the RS/IC classification accuracy. It is proved that the model in this paper not only reduces the dependence of the classification performance on the training set, but also is more robust in the classification of waveforms from different regions.

Acknowledgement

This work is supported by Science and Technology Project of State Grid Corporation of China (5500-2021120583A-0-5-SF).

Open Research

The interpretable CNN model proposed in this study and related data(Xiao et al., 2022) are available at: https://doi.org/10.5281/zenodo.7549481

References

Biagi, C. J., Cummins, K. L., Kehoe, K. E., & Krider, E. P. (2007). National lightning detection network (NLDN) performance in southern Arizona, Texas, and Oklahoma in 2003–2004. *Journal of Geophysical Research: Atmospheres*, *112*(D5). https://doi.org/10.1029/2006JD007341

Bils, J. R., Thomson, E. M., Uman, M. A., & Mackerras, D. (1988). Electric field pulses in close lightning cloud flashes. *Journal of Geophysical Research: Atmospheres*, *93*(D12), 15933–15940. https://doi.org/10.1029/JD093iD12p15933

Bitzer, P. M., Christian, H. J., Stewart, M., Burchfield, J., Podgorny, S., Corredor, D., Hall, J., Kuznetsov, E., & Franklin, V. (2013). Characterization and applications of VLF/LF source locations from lightning using the Huntsville Alabama Marx Meter Array. *Journal of Geophysical Research: Atmospheres*, *118*(8), 3120–3138. https://doi.org/10.1002/jgrd.50271

Brunner, K. N. (2016). *Explorations in intracloud lightning and leader processes*. The University of Alabama in Huntsville.

Cooray, V. (2009). Propagation Effects Due to Finitely Conducting Ground on Lightning-Generated Magnetic Fields Evaluated Using Sommerfeld's Integrals. *IEEE Transactions on Electromagnetic Compatibility*, *51*(3), 526–531. https://doi.org/10.1109/TEMC.2009.2019759

Cooray, V., & Lundquist, S. (1982). On the characteristics of some radiation fields from lightning and their possible origin in positive ground flashes. *Journal of Geophysical Research: Oceans*, *87*(C13), 11203–11214. https://doi.org/10.1029/JC087iC13p11203

Cooray, V., & Lundquist, S. (1985). Characteristics of the radiation fields from lightning in Sri Lanka in the tropics. *Journal of Geophysical Research: Atmospheres*, *90*(D4), 6099–6109. https://doi.org/10.1029/JD090iD04p06099

Da Silva, C. L., & Pasko, V. P. (2015). Physical mechanism of initial breakdown pulses and narrow bipolar events in lightning discharges. *Journal of Geophysical Research: Atmospheres*, *120*(10), 4989–5009. https://doi.org/10.1002/2015JD023209

Gomes, C., & Cooray, V. (2001). Characteristics of cloud flashes. *14th International Zurich Symposium EMC 58J3*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. https://doi.org/10.1109/cvpr.2016.90

Kašpar, P., Santolík, O., Kolmašová, I., & Farges, T. (2017). A model of preliminary breakdown pulse peak currents and their relation to the observed electric field pulses. *Geophysical Research Letters*, *44*(1), 596–603. https://doi.org/10.1002/2016gl071483

Kaur, H., Pannu, H. S., & Malhi, A. K. (2019). A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions. *ACM Comput. Surv.*, *52*(4). https://doi.org/10.1145/3343440

Kohlmann, H., Schulz, W., & Pedeboy, S. (2017). Evaluation of EUCLID IC/CG classification performance based on ground-truth data. *2017 International Symposium on Lightning Protection (XIV SIPDA)*, 35–41. https://doi.org/10.1109/SIPDA.2017.8116896

Krider, E. P., Radda, G. J., & Noggle, R. C. (1975). Regular radiation field pulses produced by intracloud lightning discharges. *Journal of Geophysical Research*, *80*(27), 3801–3804. https://doi.org/10.1029/jc080i027p03801

Le Vine, D. M. (1980). Sources of the strongest RF radiation from lightning. *Journal* *of Geophysical Research: Oceans*, *85*(C7), 4091–4095. https://doi.org/10.1029/jc085ic07p04091

Leal, A. F., Rakov, V. A., & Rocha, B. R. (2019). Compact intracloud discharges: New classification of field waveforms and identification by lightning locating systems. *Electric Power Systems Research*, *173*, 251–262. https://doi.org/10.1016/j.epsr.2019.04.016

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57.

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Lyu, F., Cummer, S. A., & McTague, L. (2015). Insights into high peak current in-cloud lightning events during thunderstorms. *Geophysical Research Letters*, *42*(16), 6836–6843. https://doi.org/10.1002/2015GL065047

Mallick, S., Rakov, V. A., Hill, J. D., Ngin, T., Gamerota, W. R., Pilkey, J. T., Jordan, D. M., Uman, M. A., Heckman, S., Sloop, C. D., & Liu, C. (2015). Performance characteristics of the ENTLN evaluated using rocket-triggered lightning data. *Electric Power Systems Research*, *118*, 15–28. https://doi.org/10.1016/j.epsr.2014.06.007

Murphy, M. J., Cramer, J. A., & Said, R. K. (2021). Recent History of Upgrades to the U.S. National Lightning Detection Network. *Journal of Atmospheric and Oceanic Technology*, *38*(3), 573–585. https://doi.org/10.1175/JTECH-D-19-

0215.1

Nag, A., Murphy, M. J., Cummins, K. L., Pifer, A. E., & Cramer, J. A. (2014). Recent evolution of the us National lightning detection network. *23rd International Lightning Detection Conference & 5th International Lightning Meteorology Conference*.

Nag, A., & Rakov, V. A. (2008). Pulse trains that are characteristic of preliminary breakdown in cloud-to-ground lightning but are not followed by return stroke pulses. *Journal of Geophysical Research: Atmospheres*, *113*(D1). https://doi.org/10.1029/2007JD008489

Nassralla, M., El Zein, Z., & Hajj, H. (2017). Classification of normal and abnormal heart sounds. *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)*, 1–4. https://doi.org/10.1109/ICABME.2017.8167538

Peng, C., Liu, F., Zhu, B., & Wang, W. (2019). A convolutional neural network for classification of lightning LF/VLF waveform. *2019 11th Asia-Pacific International Conference on Lightning (APL)*, 1–4. https://doi.org/10.1109/APL.2019.8815977

Rakov, V. A., & Uman, M. A. (2003). *Lightning: Physics and effects*. Cambridge university press.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). " Why should i trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Said, R., Inan, U., & Cummins, K. (2010). Long-range lightning geolocation using a VLF radio atmospheric waveform bank. *Journal of Geophysical Research: Atmospheres*, *115*(D23). https://doi.org/10.1029/2010JD013863

Schulz, W., Diendorfer, G., Pedeboy, S., & Poelman, D. R. (2016). The European lightning location system EUCLID –\hack\newline Part 1: Performance analysis and validation. *Natural Hazards and Earth System Sciences*, *16*(2), 595–605. https://doi.org/10.5194/nhess-16-595-2016

Shao, X.-M., & Jacobson, A. R. (2009). Model Simulation of Very Low-Frequency and Low-Frequency Lightning Signal Propagation Over Intermediate Ranges. *IEEE Transactions on Electromagnetic Compatibility*, *51*(3), 519–525. https://doi.org/10.1109/TEMC.2009.2022171

Smith, D., Shao, X., Holden, D., Rhodes, C., Brook, M., Krehbiel, P., Stanley, M., Rison, W., & Thomas, R. (1999). A distinct class of isolated intracloud lightning discharges and their associated radio emissions. *Journal of Geophysical Research: Atmospheres*, *104*(D4), 4189–4212. https://doi.org/10.1029/1998JD200045

Wang, J., Huang, Q., Ma, Q., Chang, S., He, J., Wang, H., Zhou, X., Xiao, F., & Gao, C. (2020). Classification of VLF/LF lightning signals using sensors and deep learning methods. *Sensors*, *20*(4), 1030. https://doi.org/10.3390/s20041030

Wang, Y., Gu, S., Fang, Y., Xu, Y., Chen, Y., & Li, P. (2020). Compact electric field change meter and its application in lightning detection and fault analysis for power grids. *2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE)*, 1–4. https://doi.org/10.1109/ICHVE49031.2020.9279853

Wang, Y., Qie, X., Wang, D., Liu, M., Su, D., Wang, Z., Liu, D., Wu, Z., Sun, Z., & Tian, Y. (2016). Beijing Lightning Network (BLNET) and the observation on preliminary breakdown processes. *Atmospheric Research*, *171*, 121–132. https://doi.org/10.1016/j.atmosres.2015.12.012

Wooi, C.-L., Abdul-Malek, Z., Salimi, B., Ahmad, N. A., Mehranzamir, K., & Vahabi-Mashak, S. (2015). A comparative study on the positive lightning return stroke electric fields in different meteorological conditions. *Advances in Meteorology*, *2015*. https://doi.org/10.1155/2015/307424

Wu, T., Wang, D., & Takagi, N. (2018). Locating preliminary breakdown pulses in positive cloud-to-ground lightning. *Journal of Geophysical Research: Atmospheres*, *123*(15), 7989–7998. https://doi.org/10.1029/2018JD028716

Xiao, L., Wang, Y., He, H., & Chen, W. (2022). VLF/LF lightning waveform classification version-alpha[dataset]. https://doi.org/10.5281/zenodo.7549481

Zhu, Y., Bitzer, P., Rakov, V., & Ding, Z. (2021). A machine-learning approach to classify cloud-to-ground and intracloud lightning. *Geophysical Research Letters*, *48*(1), e2020GL091148. https://doi.org/10.1029/2020GL091148