



# Application of a neural network model with multimodal fusion for fluorescence spectroscopy

Lin Tang<sup>1,2,3,4</sup> · Shuang Zhou<sup>1</sup> · Kai-Bo Shi<sup>1</sup> · Hong-Tao Shen<sup>2</sup> · Lei You<sup>5,6</sup>

Received: 27 December 2023 / Revised: 16 February 2024 / Accepted: 27 February 2024 / Published online: 25 September 2024  
© The Author(s), under exclusive licence to China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2024

## Abstract

In energy-dispersive X-ray fluorescence spectroscopy, the estimation of the pulse amplitude determines the accuracy of the spectrum measurement. The error generated by the amplitude estimation of the pulse output distorted by the measurement system leads to false peaks in the measured spectrum. To eliminate these false peaks and achieve an accurate estimation of the distorted pulse amplitude, a composite neural network model is proposed, which embeds long and short-term memory (LSTM) into the UNet structure. The UNet network realizes the fusion of pulse sequence features and the LSTM model realizes pulse amplitude estimation. The model is trained using simulated pulse datasets with different amplitudes and distortion times. For the pulse height estimation, the average relative error of the trained model on the test set was approximately 0.64%, which is 27.37% lower than that of the traditional trapezoidal shaping algorithm. Offline processing of a standard iron source further validated the pulse height estimation performance of the UNet-LSTM model. After estimating the amplitude of the distorted pulses using the model, the false peak area was reduced by approximately 91% over the full spectrum and was corrected to the characteristic peak region of interest (ROI). The corrected peak area accounted for approximately 1.32% of the characteristic peak ROI area. The results indicate that the model can accurately estimate the height of distorted pulses and has substantial corrective effects on false peaks.

**Keywords** UNet · Long- and short-term memory · Pulse distortion · Pulse height estimation · Fluorescent spectroscopy

## 1 Introduction

In a nuclear radiation measurement system, the measured spectra are often limited by distorted pulses from the detector output. Distorted pulses are non-ideal pulse signals that

occur during the reception and processing by the detector and primarily include pile-up, interference, slow, spark, double, and truncated pulses. In measurement systems that use switch reset preamplifiers, distorted pulses are mainly composed of truncated pulses. Truncated pulses are pulse signals in which the pulse amplitude suddenly jumps to zero because of a switch reset, resulting in a short pulse drop time and insufficient effective width. During the readout of nuclear pulse signals, distorted pulses output by the detector are amplified and shaped, and the height is obtained during digitization. Consequently, the measured spectra are also distorted [1]. This distortion has a significant impact on the analysis of sample element content; the count loss in the region of interest (ROI) of the target element characteristic peak leads to an underestimation of the element content.

In recent years, numerous studies have been conducted on pulse distortion in the fields of spectroscopy and radiometry. From the perspective of traditional methods, the simplest and most efficient method for generating distorted pulses is pulse elimination. We have previously detailed the method

---

This work was supported by the Open Project of Guangxi Key Laboratory of Nuclear Physics and Nuclear Technology (No. NLK2022-05), the Central Government Guidance Funds for Local Scientific and Technological Development, China (No. Guike ZY22096024), the Sichuan Natural Science Youth Fund Project (No. 2023NSFSC1366) and Key R&D Projects of Sichuan Provincial Department of Science and Technology (No. 2023YFG0287), the Open Research Fund of National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University (No. AE202209), the National Natural Science Youth Foundation of China (No. 12305214), the Vanadium and Titanium Resource Comprehensive Utilization Key Laboratory of Sichuan Province (No. 2023FTSZ03), the Key Laboratory of Interior Layout optimization and Security, Institutions of Higher Education of Sichuan Province (No. 2023SNKJ-01).

---

Extended author information available on the last page of the article

for eliminating distorted pulses and analyzed the elimination results in Ref. [2]. However, this method also has the problem with count loss. Therefore, we subsequently proposed algorithms such as signal reconstruction [3] and multipulse local spectroscopy [4] to compensate for the count loss. In addition to eliminating signal reconstruction, some scholars have proposed applying mathematical models based on the pulse shape to  $\gamma$ -ray spectroscopy [5] or applying numerical method-based filter models to neutron gamma monitoring [6], both of which are highly effective in eliminating spectral distortion.

In recent years, deep-learning technology has developed rapidly and many excellent models have emerged, such as transformer [7], UNet [8], VNet [9], and U2Net [10]. Xu et al. [11] detailed the application and development trends of artificial intelligence methods in multiple disciplines such as mathematics, materials science, medicine, life sciences, and nuclear physics. Taking nuclear physics as an example [12], various artificial intelligence methods have been widely used in advanced material calculations [13], radiation measurements [14], nuclide recognition [15], and other fields. Practical applications have also been realized, including pulse amplitude estimation [16], radiation dose measurement in the human body [17], particle-type discrimination [18], gamma spectrum analysis [19], pulse signal analysis based on feature fusion [20], and residual structure [21]. Deep-learning technology provides various ideas for pulse processing in radiation measurements. This study aimed to implement a composite neural network model that combines signal-to-noise ratio (SNR) and pulse height estimation. It is widely known that the simplest and most effective method for estimating pulse height is linear unfolding; however, this method is not immune to various types of noise. To avoid this problem, some scholars have attempted to design nonlinear filters [22]; however, the design process is more complex.

In practical applications, the commonly used method for pulse amplitude estimation is digital pulse shaping technology [23], including CR-RC filtering [24], Sallen–Key filtering, digital trapezoidal, and Gaussian shaping methods. Zhang et al. derived numerical recursive models for CR differential circuits and RC integral circuits, analyzed the amplitude–frequency response of CR-RCm, and showed that it is a bandpass filter. Wengang et al. proposed a digital trapezoidal shaping algorithm [25] and Sallen–Key filtering shaping algorithm [26]. These two shaping methods are widely used because of their relatively simple implementation, but they are easily affected by the parameter drift of the front-end detector and circuit, resulting in significant amplitude estimation errors. Gaussian shaping [27] can accurately extract the peak position and amplitude; however, this method is relatively complex and the shaping process requires precise parameter adjustments. The aforementioned

pulse amplitude estimation methods have been widely used because of their unique advantages; however, for the estimation of distorted pulse amplitudes, these methods have significant errors.

Based on the above reasons and considering that UNet has been successfully applied to audio source segmentation as a filter model [28], this study proposes an improved neural network model that uses UNet as its basic framework. The long short-term memory (LSTM) model, owing to its flexibility in handling time-series events and the information control ability of the gating mechanism, can better capture the characteristics of temporal data and make accurate predictions [29]. Therefore, this study added LSTM to the UNet framework. We successfully used the LSTM model to identify and separate stacked pulses during the early stages [30]. Accordingly, this paper presents the topology and training process of a neural network model generated using a combination of multiple models. This model was used to estimate the pulse height of an input dataset, thereby correcting the false peaks generated by the distorted pulses while ensuring the accuracy of the count rate.

The remainder of this paper is organized as follows.

Section 2 describes the generation of the simulated pulse datasets.

Section 3 describes the topological structure of the model and settings of the training parameters.

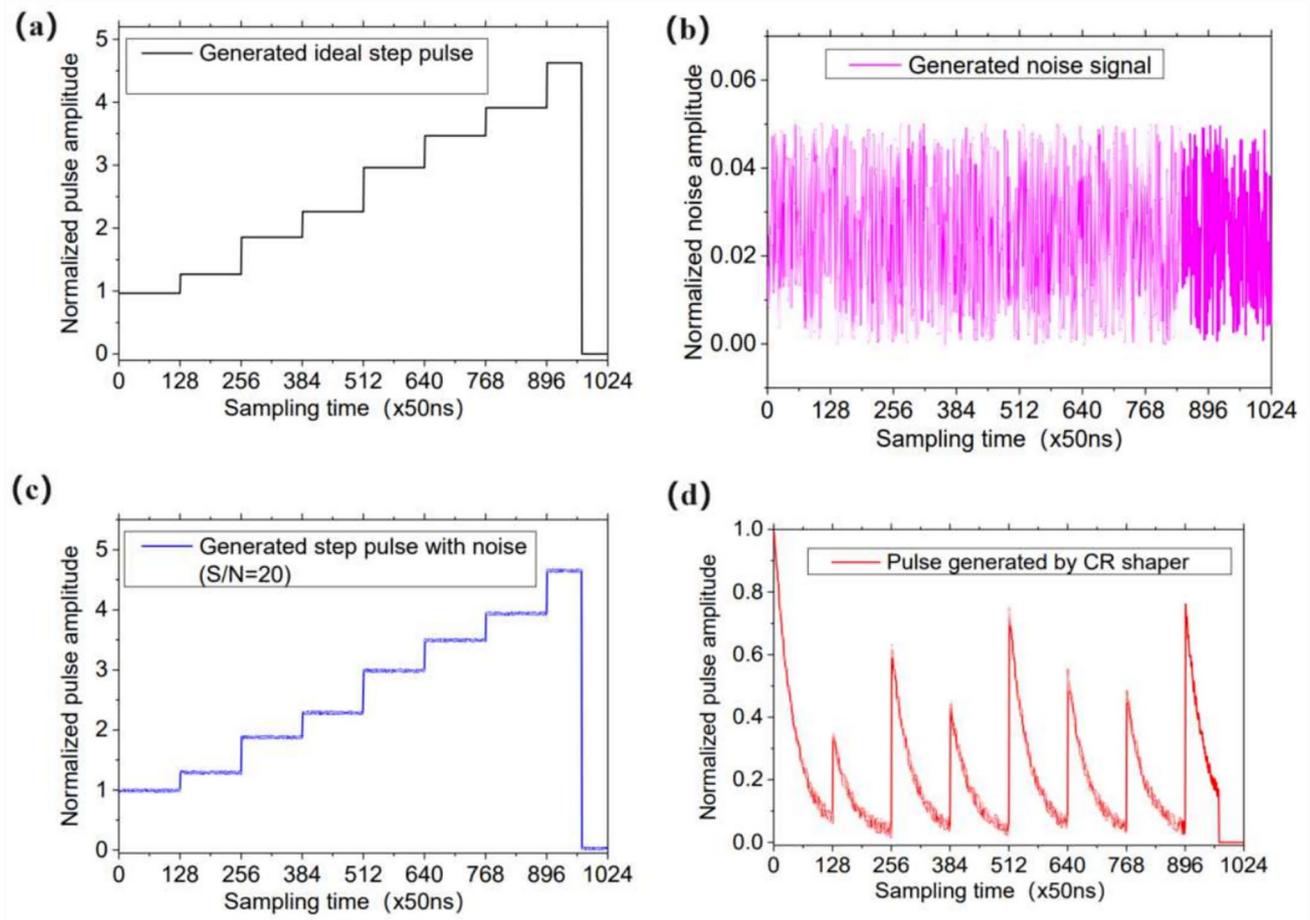
In Sect. 4, the performance of the model is evaluated using simulated and measured pulses.

Finally, Sect. 5 summarizes the conclusions of the study.

## 2 Generation of data

### 2.1 Generation of input pulses

The composite neural network model of UNet-LSTM proposed in this study realizes accurate estimations of the pulse height, especially for the height estimation of the pulse output distorted by fast silicon drift detectors (FAST SDD). The generation of the negative exponential pulse and the configuration process of the simulation links are shown in Fig. 1. One random-number generator is used to generate uniformly distributed random numbers within the 0–1 interval, which serve as the pulse heights of the stacked rising step pulses. An ideal step-pulse sequence is shown in Fig. 1a [31]. Similarly, another random-number generator is used to generate uniformly distributed random numbers within the range of 0–0.05, which are used as the amplitude of the noise signal, as shown in Fig. 1b. Random noise is added to the generated step-pulse signal to simulate a real pulse signal. The distorted pulse sequence superimposed with white noise is shown in Fig. 1c, with a SNR of 20. The stacked and rising step-pulse sequence after CR shaping is filtered



**Fig. 1** **a** Simulated ideal step pulse; **b** Noise generated by a random-number generator; **c** Step pulses with noise; **d** Pulse sequence after CR shaping

out of the DC component, and a negative exponential pulse sequence with an amplitude between 0 and 1 is output, as shown in Fig. 1d.

## 2.2 Dataset production

Traditionally, digital shaping is used for estimating the amplitude of negative exponential pulses. Taking trapezoidal shaping as an example, to achieve ideal shaping results, there need to be enough sampling points after digitizing negative exponential pulses to obtain accurate shaping results. If the pulse distortion is severe and too many sampling points are lost, the shaping results experience significant losses. The UNet-LSTM model proposed in this study is primarily aimed at estimating the height of severely distorted pulses, the characteristic of which is that the pulse height is severely damaged after trapezoidal shaping.

Nine negative exponential pulse sequences with an amplitude of 20 are shown as an example in Fig. 2a. The corresponding trapezoidal shaping results are shown in

Fig. 2b. When the negative exponential pulse loses more sampling points, the amplitude loss of the shaping result is significant, as shown by the black curve in Fig. 2b. When the number of sampling points increases to approximately 80, even after a loss of sampling information, the pulse amplitude in the shaping result is not significantly affected.

This study used two datasets: Dataset I was composed of negative exponential pulses, and Dataset II was composed of shaped triangular pulses, as shown in Fig. 2. Each dataset was divided into training, validation, and testing sets in a ratio of 7:2:1.

To train the model to estimate the amplitude of pulses with different amplitudes and degrees of distortion, the pulse amplitude was set between 20 and 1500 during dataset production, with an amplitude interval of 5. The degree of distortion of the negative exponential pulses is determined by the number of sampling points. The number of sampling points was set between 40 and 80 with an interval of 1, resulting in a distorted pulse dataset with a

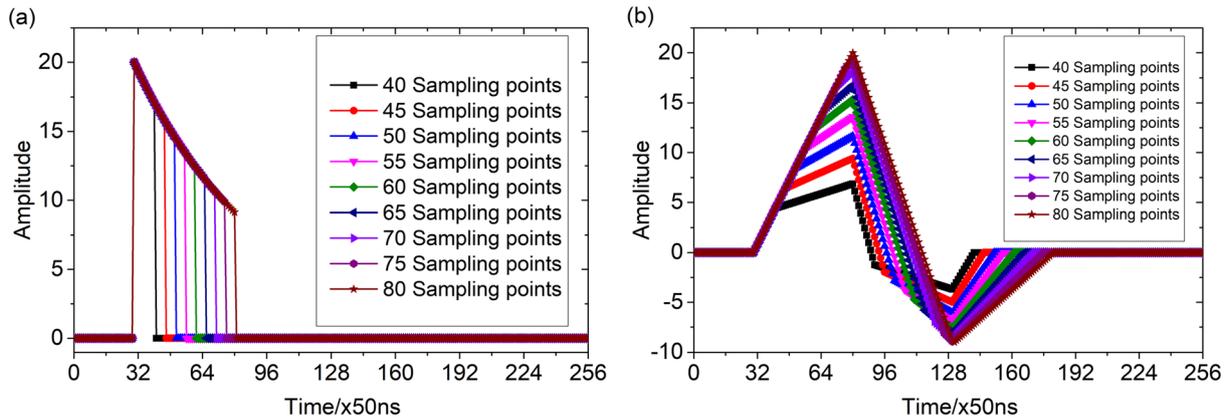


Fig. 2 (Color online) a Negative exponential pulse sequence; b Digital shaping result

Table 1 Dataset details

Description	Size
Amplitude range	20–1500 (interval is 5)
Number of pulses per amplitude	41
Sampling point range	41–80 (interval is 1)
Dataset size	121770×256
Parameter set size	121770×2
Training set size	85239×256
Validation set size	24354×256
Test set size	12177×256

size of  $121770 \times 256$ . Detailed information regarding the dataset is presented in Table 1.

Dataset I is taken from the pulse amplitude sampling value of the distortion pulses after the CR shaper, and the sampling period is  $T_s$ , whereas the parameter set  $P$  is taken from the amplitudes of the pulses and the number of sampling points. The matrix representation of the dataset is expressed by Eq. (1).

$$\begin{pmatrix} [V(T_s)]_1 & [V(2T_s)]_1 & \cdots & [V(nT_s)]_1 & P_1 \\ [V(T_s)]_2 & [V(2T_s)]_2 & \cdots & [V(nT_s)]_2 & P_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ [V(T_s)]_N & [V(2T_s)]_N & \cdots & [V(nT_s)]_N & P_N \end{pmatrix} \quad (1)$$

Dataset I, as shown in Eq. (1), contains  $N = 121770$  pulses. Each pulse corresponds to a row in the matrix, and each row contains  $n + 1$  columns. The first  $n$  columns correspond to the sampling values of each pulse and the last column represents the parameters. As required, this study divided the dataset into a training set, a validation set, and a test set according to a ratio of 7:2:1, ensuring that they were divided according to this ratio for each amplitude interval. Figure 3 presents the details of the two datasets. Figure 3a shows eight negative exponential pulses and their matrix

expressions obtained from Dataset I. Figure 3b shows the results obtained by shaping the eight negative exponential pulses and their matrix expressions in Dataset II. The essential difference between them is that Dataset II is obtained by pulse shaping Dataset I and maintains the same data size.

### 3 Model development

Convolutional neural networks (CNNs) are currently the most popular deep-learning methods. The biggest difference from a fully convolutional network (FCN) is that the neurons in each layer of a CNN focus only on a certain part of the signal to analyze a specific feature and are not connected to all the neurons in the next layer. Compared to FCN, it saves resources and training time.

The UNet model is a CNN model used for image segmentation. It has two parts, an encoder and a decoder, which restore the detailed information of the image by connecting low- and high-level features. The model adopts a U-shaped connection structure that can effectively handle the detailed information in the input object. The LSTM model is a recursive neural network model used for processing sequence data. It has a gating mechanism that can selectively remember and forget previous information and capture long-term dependencies, making it suitable for processing time-series data. In this study, UNet and LSTM models were combined to form a composite neural network model. When selecting model parameters, two principles are followed. First, classical model parameters are referenced for UNet [8], and then local adjustments are made to the LSTM model based on the specific tasks and data characteristics.

The classic UNet model includes an encoder for extracting signal features on the left and a decoder for feature fusion on the right. Both the encoder and decoder contain four sets of two  $3 \times 3$  convolutional layers. The difference is that the encoder uses a maximum pooling operation for

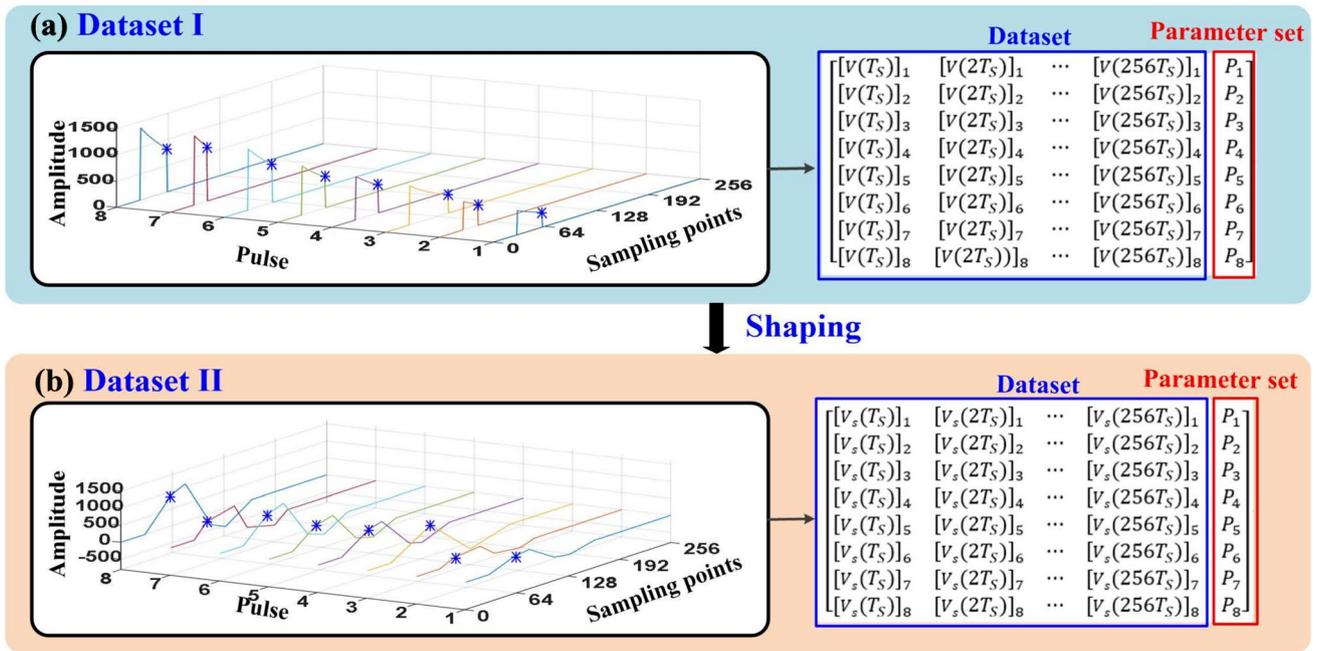


Fig. 3 (Color online) **a** Negative exponential pulse sequence and its matrix; **b** Digital shaping result and its matrix

downsampling after the convolutional layer, with a pooling size of  $2 \times 2$  and a step size of 2, whereas the decoder first upsamples the input signal through  $2 \times 2$  transposed convolution operations to reduce the number of feature channels, and then performs a convolution operation.

Because of the recursive structure of LSTM, it is difficult to achieve parallel operations, and gradient dispersion may occur when the sequence is very long. In this study, the input sequence length was limited for the amplitude parameter estimation task. Given the powerful feature-extraction ability of LSTM, it typically does not require many layers. The number of layers in our LSTM was set to 1.

During the model training process, the selection of the activation and loss functions followed the principle of local adjustment according to the characteristics of the task in the second item. The aim of this study was to estimate the pulse amplitude parameters, and the data type is one-dimensional linear time-series pulses; therefore, the activation function used ReLU (Rectified Linear Unit). The input and output of the model are both pulse amplitude information. Therefore, the loss function uses the mean square error (MSE).

Figure 4 illustrates the internal structure of the proposed composite neural network model. The model was trained using the dataset described in Sect. 2. Among the 121770

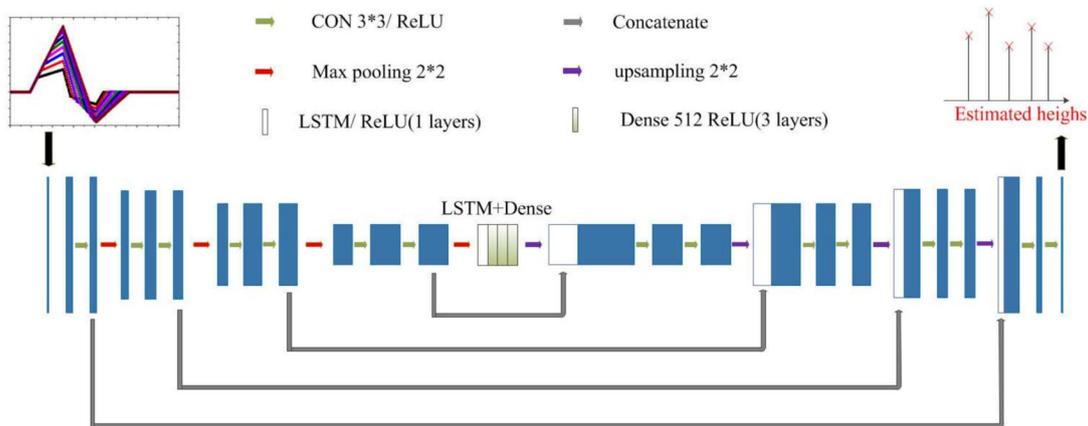


Fig. 4 (Color online) UNet-LSTM model structure

pulses in the dataset, 85239 pulses were used as the training set, 24354 pulses were used as the validation set, and 12177 pulses were used as the testing set.

In the training process of the UNet-LSTM model, the input signal is a distorted pulse superimposed with white noise, and the output signal is a set of expanded pulse heights. The model was trained using the Adam optimizer. The learning rate is an important hyperparameter in model training and was defined as the cyclical learning rate (CLR) in the range from  $1 \times 10^{-5}$  to  $1 \times 10^{-3}$ . To evaluate the difference between the model output and expected output, the MSE loss function  $L_{\text{MSE}}$  provided feedback to the network to update the weights and reducing subsequent iteration errors. The error between the model output pulse height set  $P'_i$  and expected output value  $P_i$  can be calculated using the loss function. For a training set with  $N$  samples, the calculation for the loss function is given by Eq. (2). The model training results are elaborated in detail in the next section.

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (P_i - P'_i)^2 \quad (2)$$

## 4 Model performance evaluation

### 4.1 Simulation research

The UNet-LSTM model was trained using the Adam Optimizer at a fixed learning rate. All experiments were conducted using Pytorch 1.7, Python 3.7, and an Ubuntu 18.04 system with a 3.70 GHz i7-8700K CPU and a 32 G V100 GPU.

#### 4.1.1 Ablation study

The purpose of the ablation experiment was to study the impact of removing specific parts of the composite model on its overall performance. A lack of performance loss after removing some parts indicates that these parts are less important in the composite model. In contrast, if the performance of the model decreases significantly after removal, the design of the parts is considered essential. To conduct

ablation research on the proposed deep-learning model, three deep-learning models were implemented: LSTM, UNet, and UNet-LSTM, and their amplitude prediction performances were evaluated and compared. To control the influence of the parameters on the model performance, the composite and single models used the same parameters during the ablation study. The main parameters of the model and training results are listed in Table 2. Although the independent LSTM and UNet models can converge normally, the proposed composite model achieved lower training and validation losses under the same parameter configuration; after fusing the two models, better pulse amplitude prediction performance was achieved.

#### 4.1.2 Evaluation of parameter estimation performance

Existing pulse amplitude calculation methods are mostly digital shaping methods, which are effective for estimating the amplitude of most nuclear pulses and are therefore widely used in digital energy spectrometers. References [2, 3] indicate that the digital shaping method poses significant challenges in the amplitude estimation of pulses with shape distortion. This study further explains the relationship between the amplitude estimation results of existing methods and the degree of pulse distortion, and it demonstrates the effectiveness of the UNet-LSTM model in addressing the challenges in distorted pulse height estimation.

Twenty pulses in the test set and the amplitude prediction output of the model were subjected to inverse normalization processing. Twenty pulses with different degrees of pulse distortion are shown in Fig. 5 with the ground truth fixed at 600.

Each distortion pulse in the test set contained different effective sampling points representing varying degrees of distortion. The pulse height estimation results obtained using the UNet-LSTM model and the digital trapezoidal shaping method are shown in Fig. 6. As the degree of pulse distortion decreased, the error in the pulse heights obtained using the digital shaping method also gradually decreased. For example,  $P_1$  contains 40 valid sampling points representing the pulse with the most severe distortion in the test set. A maximum relative error of 65.67% was obtained when the digital trapezoidal shaping method

**Table 2** Ablation study comparison

Experimental subjects		Batch size	Learning rate	Train loss	Validation loss	Numbers of parameters (M)
Single model	LSTM	1	0.0002	160.8187833	69.3094321	1.97
	UNet	1	0.0002	94.6813544	20.3280027	2.11
Composite model	UNet-LSTM	1	0.0002	74.232632	10.111189	3.29

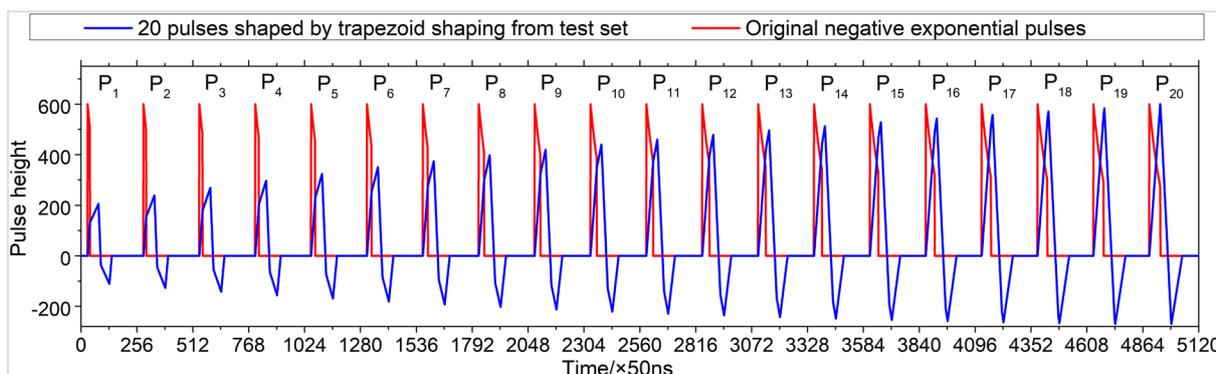


Fig. 5 (Color online) Simulated pulse sequence diagram

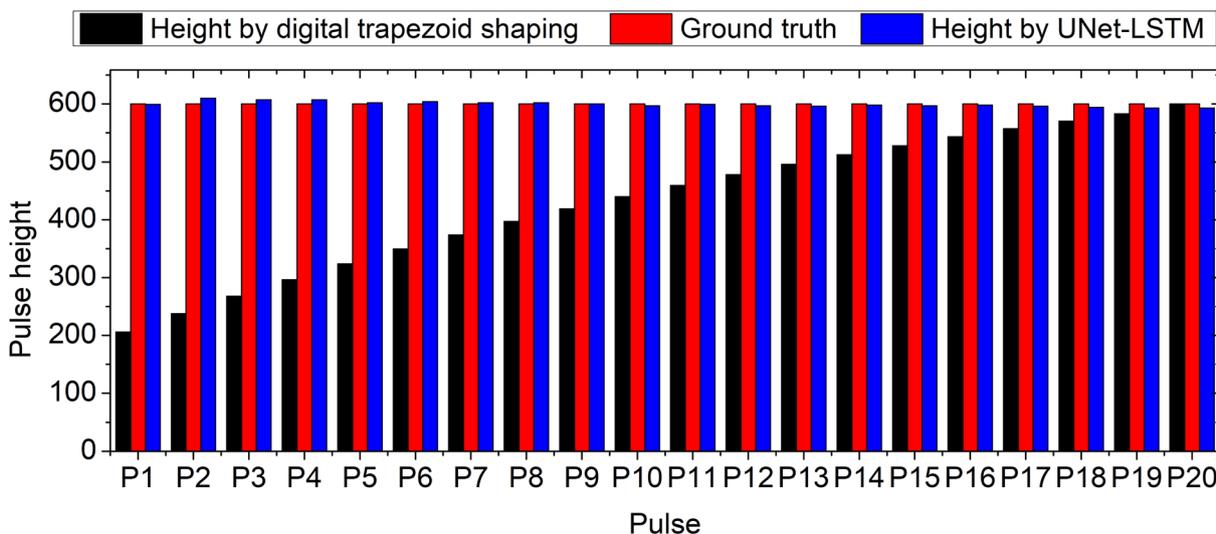


Fig. 6 (Color online) Comparison of pulse height estimation

was used to estimate the amplitude. As the number of sampling points increased, the degree of pulse distortion decreased, and the relative error of the digital trapezoidal shaping gradually decreased. In Fig. 5, the rise time of digital trapezoidal shaping includes 80 sampling points, and the flat top time is 0. Therefore, when the number of sampling points for the original negative exponential pulse was 80, the relative error of the corresponding digital shaping result was reduced to 0.02%.

To quantify the prediction ability of the UNet-LSTM model for pulse height,  $\Delta_A$  represents the absolute error of the pulse height estimation,  $\delta_A$  represents the relative error,  $A_{real}$  represents the true pulse height, and  $A_{NN}$  represents the pulse height predicted by the neural network model, UNet-LSTM. Thus, the calculations for the absolute and relative errors of the UNet-LSTM model for the distorted pulse height estimation are shown in Eq. (3) and Eq. (4), respectively.

$$\Delta_A = ABS(A_{real} - A_{NN}) \tag{3}$$

$$\delta_A = \frac{\Delta_A}{A_{real}} \times 100\% \tag{4}$$

The 20 pulses shown in Fig. 6 were analyzed for pulse height using the traditional trapezoidal shaping algorithm and the neural network model proposed in this study. The results are shown in Table 3. The average relative error of the trapezoidal shaping algorithm in estimating the pulse height of the sequence is 28.01%, while that of the UNet-LSTM model is approximately 0.64%. This proves that the UNet-LSTM model is not affected by pulse distortion when estimating pulse amplitude parameters, addressing the limitations facing existing methods in pulse height estimation.

Each distorted pulse in the test set contains different effective sampling points representing different degrees of

**Table 3** Comparison between the estimated and true values of pulse height using neural network models

Pulse	Real height	Pulse height by trapezoidal shaping	Relative error of trapezoidal shaping (%)	Estimated value by model	Relative error of the UNet-LSTM model (%)
P1	600.000	205.989	65.67	599.000	0.17
P2	600.000	237.709	60.38	610.000	1.67
P3	600.000	267.851	55.36	607.000	1.17
P4	600.000	296.477	50.59	607.000	1.17
P5	600.000	323.656	46.06	602.000	0.33
P6	600.000	349.448	41.76	604.000	0.67
P7	600.000	373.913	37.68	602.000	0.33
P8	600.000	397.109	33.82	602.000	0.33
P9	600.000	419.093	30.15	600.000	0.00
P10	600.000	439.911	26.68	597.000	0.50
P11	600.000	459.618	23.40	599.000	0.17
P12	600.000	478.265	20.29	597.000	0.50
P13	600.000	495.895	17.35	596.000	0.67
P14	600.000	512.555	14.57	598.000	0.33
P15	600.000	528.287	11.95	597.000	0.50
P16	600.000	543.137	9.48	598.000	0.33
P17	600.000	557.133	7.14	596.000	0.67
P18	600.000	570.324	4.95	594.000	1.00
P19	600.000	582.749	2.88	593.000	1.17
P20	600.000	600.121	0.02	593.000	1.17

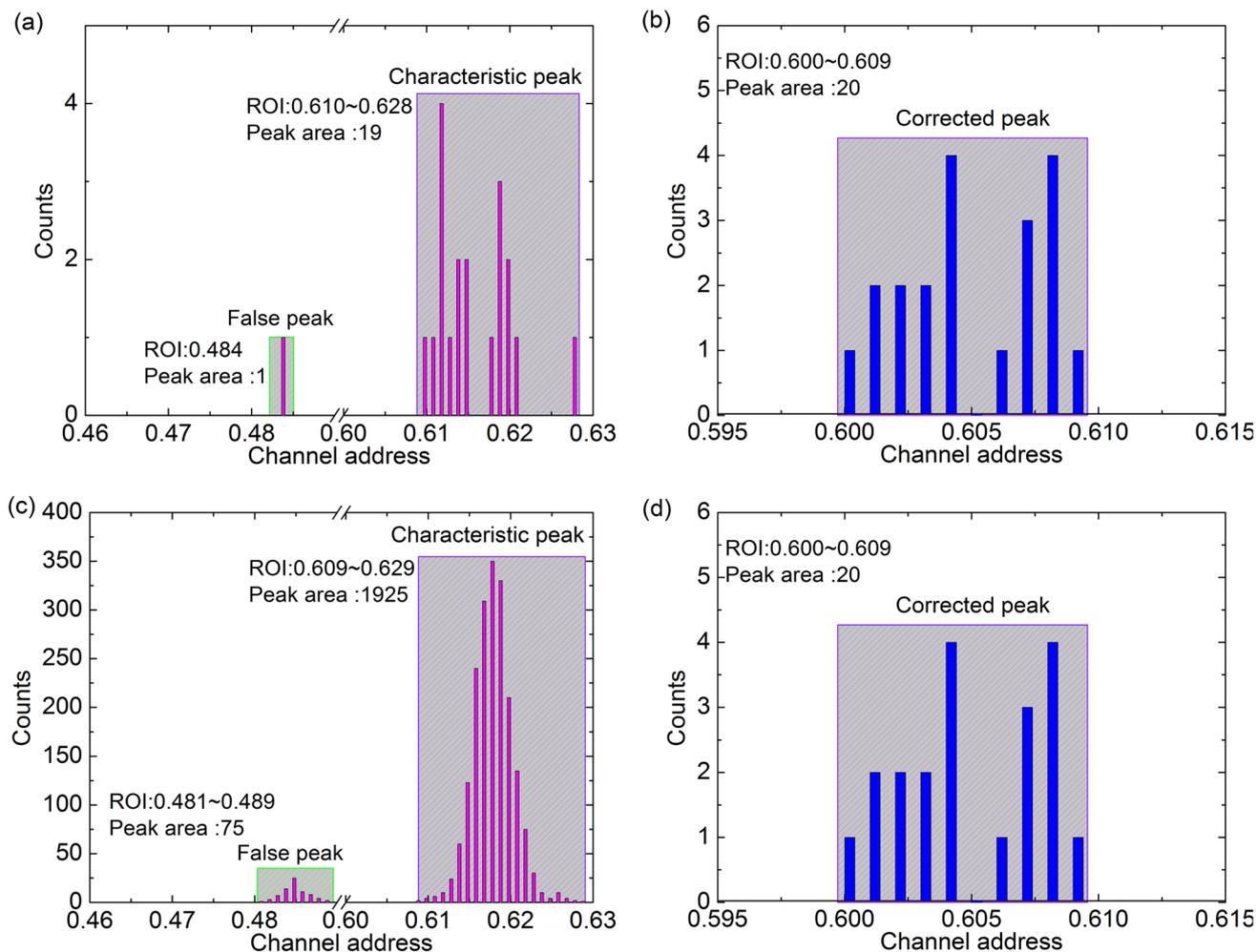
distortion. For example, P1 contains 40 sampling points, representing the most severely distorted pulse in the test set, and its corresponding digital shaping results had a maximum relative error of 65.67%. As the number of sampling points increased, the degree of pulse distortion decreased, and the relative error of the digital trapezoidal shaping gradually decreased. In Fig. 5, the rise time of digital trapezoidal shaping includes 80 sampling points, and the flat top time is 0. Therefore, when the number of sampling points for the original negative exponential pulse was 80, the relative error of the corresponding digital shaping results was reduced to 0.02%. Finally, the average relative error of the trapezoidal shaping algorithm for estimating the pulse height in this sequence was 28.01%. Using the UNet-LSTM model to unfold the input pulse sequence and estimate the pulse height, the estimations are not affected by pulse distortion, with an average relative error of approximately 0.64%.

#### 4.1.3 Evaluation of count correction performance

According to the principle of a multichannel analyzer (MCA), each pulse amplitude in the X-ray fluorescence spectrum corresponds to a count in the energy histogram. When the pulse output of the measurement system experiences an amplitude loss during the digital processing stage, the corresponding histogram drifts to the left. When the number of pulses is sufficient, false characteristic peaks exist in the

generated energy spectrum. For samples with a single component, the number of characteristic peaks of the elements in the measured full spectrum is limited. Therefore, the pulse amplitude output of the measurement system mostly fluctuates within a small range of fixed values, which determine the energy resolution of the spectrum. Quantification was performed based on the half-widths of the characteristic peaks in the spectrum. The 20 simulated pulses mentioned earlier were analyzed using traditional MCA as an example. Owing to the normalization of the input pulse amplitude in the model, the channel address range of the full spectrum was set to 0~1, including 1000 channel addresses. When the pulse with an amplitude of 0.6 is increased by one, the count on the channel address of 0.6 increases by one, resulting in an X-ray spectrum, as shown in Fig. 7a.

In this figure, the channel address range of the characteristic peak ROI is approximately 0.610–0.628 with a peak area of 19. A false peak formed by distorted pulses appeared near the 0.484th channel address on the left side of the characteristic peak, with a peak area of 1, as indicated by the green shaded area in Fig. 7a. The CNN-LSTM model proposed in this study was added before the MCA unit to achieve an accurate estimation of the pulse parameters. A histogram of the characteristic peaks obtained using the model is shown in Fig. 7b, with the ROI of the characteristic peaks reduced to approximately 0.600–0.609; however, the total count within this range increased to 20.



**Fig. 7** (Color online) Simulated energy spectrum **a** without the model and with a total count of 20; **b** with the model and a total count of 20; **c** without the model and with a total count of 2000; **d** with the model and a total count of 2000

Therefore, using the model to estimate the pulse amplitude not only ensures that the counting of characteristic peak areas is not lost but also eliminates false peaks caused by pulse distortion.

## 4.2 Experimental results

### 4.2.1 Analysis of generalizability experiment results

To conduct a more thorough statistical analysis of the model performance and address potential biases in the simulation dataset used for model training, a <sup>55</sup>Fe standard source was used as a test object in the experimental testing phase to provide sufficient pulse sequences for the analysis of the experimental results. The experimental conditions were configured as listed in Table 4.

Using the experimental platform listed in Table 4, the pulse sequence during the measurement process was

**Table 4** Details of the experimental setup

Category	Details
Source	KYW2000A X-ray tube with Ag target Current: 8 μA Maintain voltage: 35 kV
Samples	<sup>55</sup> Fe standard source (Count rate: 7.553×10 <sup>3</sup> cps)
Detector	FAST SDD (123 eV FWHM Resolution @ 5.9 keV) Application area: high counting rate>1,000,000 CPS Be window (0.5 mil) Footprint: TO-8 Active area: 50 mm <sup>2</sup>
Digital system	ADC9235 with a resolution of 12 bits Sampling frequency: 20 Msps

**Table 5** Comparison of different models

Model	Dataset II (Simulated)		Dataset III (Measured)	
	Train loss	Validation loss	Validation loss	Numbers of parameters
LeNet5	138.69	40.29	86.18	0.13M
CNN-LSTM	242.33	28.72	68.76	1.10M
LSTM	160.82	69.30	96.11	1.97M
UNet	96.50	18.84	26.54	2.11M
UNet-LSTM	74.23	10.11	23.14	3.29M

saved offline to obtain the measured pulse dataset. The pulse dataset was processed according to the processing method of Dataset II to obtain a negative exponential pulse dataset with approximately 40–80 effective sampling points. After trapezoidal shaping of the pulses in this dataset, the required validation dataset was obtained, which was defined as Dataset III, and used to validate the trained model and demonstrate its generalizability. Dataset III contained 5000 pulses with an amplitude range of approximately 20–2000 mv. Each pulse contained 256 sampling points and one amplitude parameter.

The proposed model was compared with other state-of-the-art models to perform pulse height estimation tasks on simulated and measured datasets. The results are shown in Table 5. The selected control model includes both lightweight single models such as LeNet and composite models such as CNN-LSTM, which have been demonstrated to be effective for pulse estimation [30]. The comparison results in Table 5 further demonstrate that because the measured pulse contains more uncertainty than the simulated pulse, the validation loss of each model on the measured pulse dataset increases compared to the simulated dataset. However, UNet-LSTM still performed better than the other models on different datasets, demonstrating good robustness and generalizability.

It is worth noting that although UNet-LSTM achieved better performance on two different datasets, its parameter count was also much larger than that of the other models. Models like LeNet with small parameter numbers and short computation times, can also perform well in most simple tasks after training. Therefore, in application scenarios that require small, lightweight models, UNet-LSTM is not the best choice. This composite model with many parameters and complex calculations is more suitable for more complex time-series analyses, such as the nuclear pulse height estimation and fluorescence spectroscopy analysis tasks in this study.

**Table 6** Techniques for optimizing fluorescence spectra

Technique type	Object	Example
Category I	Nuclear pulse	Digital trapezoidal shaping
Category II	Fluorescence spectra	Spectra smooth

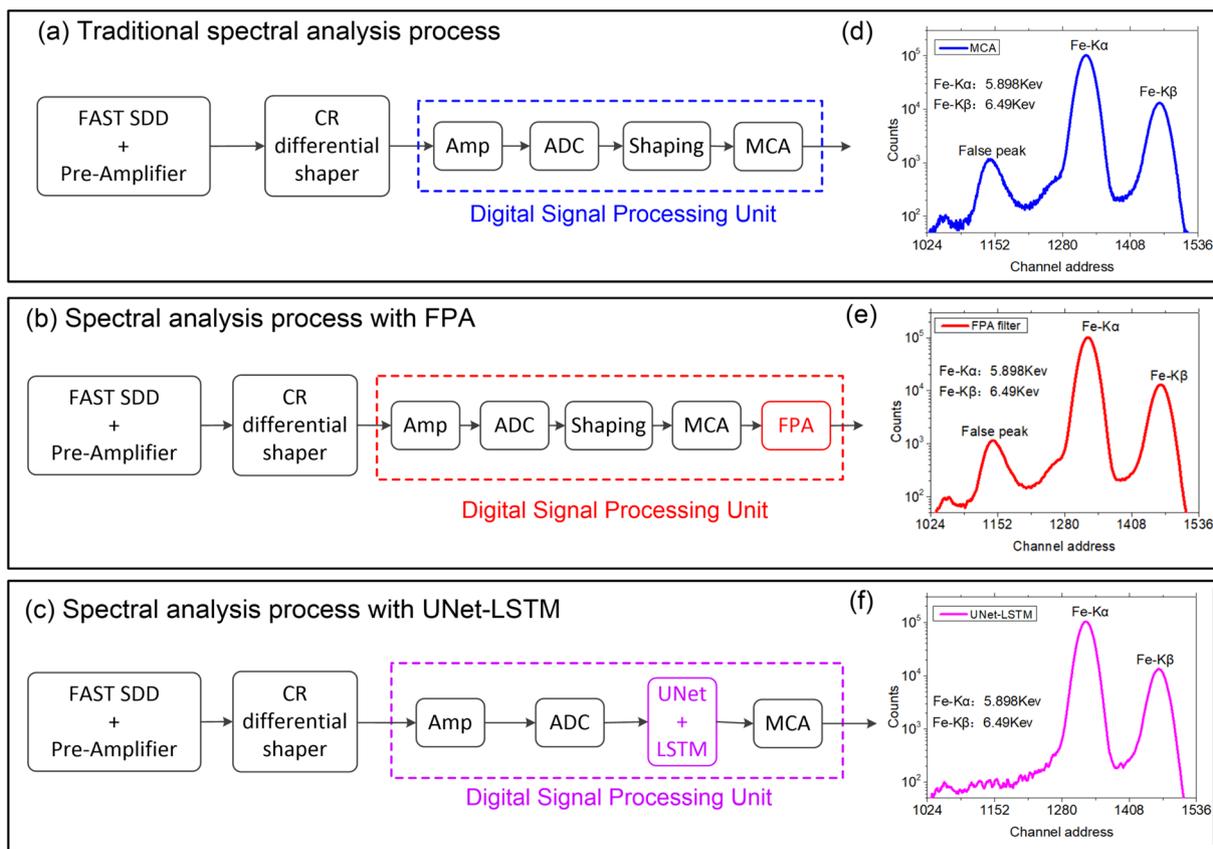
#### 4.2.2 Analysis of spectral experiment results

The commonly used techniques for optimizing fluorescence spectra can be divided into two categories from the perspective of processing objects, as shown in Table 6.

Category I is for nuclear pulses with the main purpose of obtaining more reliable pulse heights. Category II is for the spectrum itself, of which the most representative and widely used is spectral smoothing. In the simulation research section, a comparison was made between the performance of the traditional digital trapezoidal shaping method and the classical neural network model in pulse height estimation tasks, all of which belong to Category I spectral processing techniques.

Here, the experimental conditions listed in Table 4 were used to analyze the results of the spectral optimization experiments. The energy spectrum obtained from the digital MCA was used as the reference spectrum, and on this basis, the spectral smoothing and pulse height estimation units were added separately. The spectral processing flow and the results are shown in Fig. 8. The traditional spectrum acquisition process primarily includes a probe section composed of a high-performance silicon drift detector (FAST SDD) and preamplifier, a CR differential shaping section, and a digital signal processing unit composed of an operational amplifier, high-precision ADC, digital shaper, and MCA, as shown in Fig. 8a. The difference in the spectral analysis process between the added spectral smoothing units and traditional methods lies in the digital signal processing. Although traditional spectral analysis also has functions, such as shaping and filtering in digital signal processing, it does not add a dedicated spectral smoothing unit located after the MCA. In this study, a five-point average (FPA) was used to smooth and filter the generated spectrum, as shown in Fig. 8b. In another control group, the trained UNet-LSTM model was used to estimate the height of the digitized pulse sequence, and the modified energy spectrum was obtained through an MCA. The spectrum processing is shown in Fig. 8c.

The distorted pulse obtained during the measurement process was amplified and trapezoidal to obtain a distorted pulse height. This distortion in the pulse height exists in the form of a false peak before the characteristic peak in the spectrum obtained by traditional spectroscopic methods, as shown in Fig. 8d, e. Although the spectral smoothing method using multipoint averaging can optimize the spectrum to a certain extent, it has no effect on the false peaks caused by distorted pulses.



**Fig. 8** Spectral analysis process with **a** the traditional method, **b** FPA, and **c** UNet-LSTM. The energy spectrum of **d** the traditional method, **e** FPA, and **f** UNet-LSTM

When using the UNet-LSTM model to estimate the measured pulse height, the model can accurately output the pulse height with high accuracy, even for distorted pulses with incomplete widths. Therefore, after predicting the pulse height using the model, the energy spectrum obtained by the MCA can effectively correct the false peaks caused by pulse distortion, as shown in Fig. 8f.

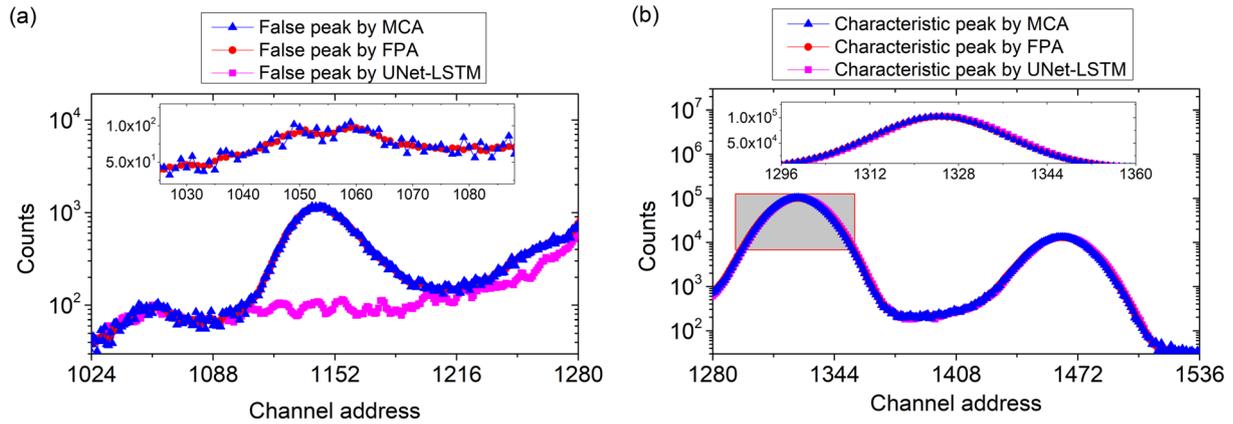
To quantify the corrective effect of the UNet-LSTM model on the measured X-ray spectrum of the <sup>55</sup>Fe standard source, two indicators, the correction ratio  $R_{correct}$  and the effective ratio  $R_{effect}$ , are introduced.  $R_{correct}$  represents the proportion of the difference in the peak area of the false peak, before and after calling the model to the peak area of the characteristic peak.  $R_{effect}$  represents the proportion of the increase in the characteristic peak area after calling the model to the loss value of the false peak area. The calculations are given by Eq. (5) and Eq. (6).

$$R_{correct} = \frac{S_{false-MCA} - S_{false-UNet}}{S_{ROI-MCA}} \times 100\% \tag{5}$$

$$R_{effect} = \frac{S_{ROI-UNet} - S_{ROI-MCA}}{S_{false-MCA} - S_{false-UNet}} \times 100\% \tag{6}$$

As shown in Fig. 9, the false peak area was located in the channel interval of approximately 1024–1280, and the ROI of the characteristic peak was captured in the channel interval of approximately 1280–1536. The peak area is represented by  $S$ ,  $S_{false-MCA}$  represents the peak area of the false peak in the traditional spectral analysis results,  $S_{false-UNet}$  represents the peak area of the false peak after calling the UNet-LSTM model for spectral analysis correction,  $S_{ROI-MCA}$  represents the peak area of the characteristic peak ROI in the traditional spectral analysis results, and  $S_{ROI-UNet}$  represents the peak area of the characteristic peak ROI after calling the UNet-LSTM model for spectral correction.

Ten measurements of the <sup>55</sup>Fe standard source were taken for the analysis, as shown in Table 5. The comparison shows that the peak area of the characteristic peak ROI in the energy spectrum obtained using the traditional MCA was similar to that obtained using FPA filtering. However, the filtered energy spectrum had smoother spectral lines in the low count-rate region; therefore, it has a lower standard deviation than traditional MCA methods in multiple measurement



**Fig. 9** (Color online) Peak area analysis of **a** a false peak and **b** characteristic peak

processes. Although FPA filtering can reduce the standard deviation of the measurement and stabilize the measurement results, it has no corrective effect on false peaks in the X-ray spectrum. The <sup>55</sup>Fe standard source X-ray spectrum obtained using the UNet-LSTM model to predict the pulse height has two typical features. First, the peak area of the characteristic peak ROI was improved, and the standard deviation of multiple measurement results was significantly reduced. Second, the peak area of false peaks was significantly reduced. According to the above two characteristics, combined with the theorem of energy conservation, it can be inferred that the reduced false peak area should be corrected to the characteristic peak ROI, and the correction effect can be evaluated using the  $R_{\text{valid}}$  index defined above. From Table 5, it can be observed that approximately 91% of the peak area loss in the false peak area can be corrected to the characteristic peak ROI, and the proportion of the corrected

peak area to the peak area of the characteristic peak ROI was approximately 1.32%, which is an indispensable part of high-count-rate applications (Table 7).

### 5 Conclusion

In this study, a composite neural network model based on the UNet architecture and fused with LSTM was proposed to achieve accurate pulse height estimations of distorted pulse sequences and thus achieve counting rate correction of the X-ray energy spectrum. The UNet part of this composite model includes an encoder for extracting pulse features and a decoder for feature fusion. Both the encoder and decoder contain eight  $3 \times 3$  convolutional layers connected by an LSTM between the encoder and decoder. The UNet-LSTM model was trained with the pulse sequence

**Table 7** Details of measurement results

	MCA		FPA		UNet-LSTM		$R_{\text{correct}}(\%)$	$R_{\text{effect}}(\%)$
	$S_{\text{ROI-MCA}}$	$S_{\text{false-MCA}}$	$S_{\text{ROI-FPA}}$	$S_{\text{false-FPA}}$	$S_{\text{ROI-UNet}}$	$S_{\text{false-UNet}}$		
1	3656332	75575	3656296	75639	3700286	27508	1.31	91.44
2	3654968	75678	3657689	75699	3699357	27488	1.32	92.11
3	3655993	75487	3654376	75589	3701056	27493	1.31	93.89
4	3654367	75651	3654789	75514	3701009	27515	1.32	96.90
5	3657547	75698	3657854	75593	3699689	27528	1.32	87.49
6	3657256	75587	3656543	75532	3699278	27541	1.31	87.46
7	3654312	75712	3654897	75710	3700179	27533	1.32	95.20
8	3658565	75613	3656451	75721	3700791	27498	1.32	87.76
9	3656843	75666	3656120	75709	3700167	27521	1.32	89.99
10	3657963	75599	3655769	75619	3700665	27509	1.31	88.80
Average	3656415	75626.6	3656078.4	75632.5	3700247.7	27513.4	1.32	91.10
STD	1417.26	64.69	1104.66	72.064	613.25	16.63	–	–

datasets generated by the simulation, and the optimal training parameters were saved when the minimum loss value was achieved for both the training and validation sets. The model performance was verified using simulated and measured pulses.

During the verification process of the simulated pulses, this study took 20 distorted pulses from the test set for pulse height estimation and the average relative error of the trained model on the test set was approximately 0.64%, which was 27.37% lower than that of the traditional trapzoidal shaping algorithm.

During the experiment, a FAST SDD was used to perform the X-ray measurements on a  $^{55}\text{Fe}$  standard source. The measured pulse sequence was saved offline as the model input, and the pulse amplitude output of the model was analyzed for an X-ray energy spectrum with the correction of false peaks. Simultaneously, the traditional MCA spectrum and a spectrum with FPA filtering were used as reference spectra for comparison with the corrected spectrum. The results indicate that the model successfully predicts the height of the measured pulse sequence. In the qualitative analysis of the  $^{55}\text{Fe}$  standard source, the spectral comparison results obtained by the three different methods indicate that although FPA filtering can achieve spectral smoothing, it has no substantial impact on false peaks, whereas the UNet-LSTM model can effectively correct false peaks caused by distorted pulses. To further validate the performance of false peak correction, the correction ratio and effective ratio were defined as new indicators of model performance. Ten measurements of the  $^{55}\text{Fe}$  standard source showed that approximately 91.1% of the false peak area could be corrected to the characteristic peak ROI, and the proportion of the corrected peak area to that of the characteristic peak ROI was approximately 1.32%. This is of great significance to optimizing X-ray energy spectrum analysis.

The neural network model proposed in this study is applicable to a wider range of detectors. In future research, we will focus on the application of this model to fast spectroscopy and improve the analysis performance of spectroscopy by accurately predicting the pulse height, which is of great significance for spectral refinement and element content analyses.

**Author contributions** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Lin Tang, Shuang Zhou, Kai-Bo Shi, Hong-Tao Shen and Lei You. The first draft of the manuscript was written by Lin Tang, and all authors commented on previous versions of the manuscript. All authors read and approved the manuscript.

**Data availability** The data that support the findings of this study are openly available in Science Data Bank at <https://doi.org/10.57760/sciencedb.j00186.00273> and <https://cstr.cn/31253.11.sciencedb.j00186.00273>.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. D. Lee, K. Lim, K. Park et al., An innovative method to reduce count loss from pulse pile-up in a photon-counting pixel for high flux X-ray applications. *J. Instrum.* **12**, P03006 (2017). <https://doi.org/10.1088/1748-0221/12/03/P03006>
2. L. Tang, J. Yu, J.B. Zhou et al., A new method for removing false peaks to obtain a precise X-ray spectrum. *Appl. Radiat. Isot.* **135**, 171–176 (2018). <https://doi.org/10.1016/j.apradiso.2018.01.033>
3. L. Tang, J.B. Zhou, F. Fang et al., Counting-loss correction for X-ray spectra using the pulse-repairing method. *J. Synchrotron Radiat.* **25**, 1760–1767 (2018). <https://doi.org/10.1107/S160057751801411X>
4. T. Lin, W.D. Zhao, S.K. Yu et al., Optimization design of X-ray spectrum data processing platform. *Spectrosc. Spectr. Anal. (in Chinese)* **41**, 763–767 (2021). [https://doi.org/10.3964/j.issn.1000-0593\(2021\)03-0763-05](https://doi.org/10.3964/j.issn.1000-0593(2021)03-0763-05)
5. M. Lee, D. Lee, E. Ko et al., Pulse pileup correction method for gamma-ray spectroscopy in high radiation fields. *Nucl. Eng. Technol.* **52**, 1029–1035 (2020). <https://doi.org/10.1016/j.net.2019.12.003>
6. B. Liu, M. Liu, M. He et al., Model-based pileup events correction via Kalman-filter tunnels. *IEEE Trans. Nucl. Sci.* **66**, 528–535 (2018). <https://doi.org/10.1109/TNS.2018.2885074>
7. L. Tang, K. Shi, H. Shen et al., Application of transformer model in peak correction of X-ray fluorescence spectra. *IEEE Trans. Nucl. Sci.* **70**, 2479–2489 (2023). <https://doi.org/10.1109/TNS.2023.3320807>
8. O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in *Medical image computing and computer-assisted intervention-MICCAI 2015 - LNCS*, vol. 9351, ed. by N. Navab, J. Hornegger, W. Wells, A. Frangi (Springer, Cham, 2015)
9. F. Milletari, N. Navab, S.A.V. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, in 2016 Fourth International Conference on 3D Vision (3DV) (pp. 565–571). IEEE
10. X. Qin, Z. Zhang, C. Huang et al., U2-net: going deeper with nested U-structure for salient object detection. *Pattern Recognit.* **106**, 107404 (2020). <https://doi.org/10.1016/j.patcog.2020.107404>
11. Y. Xu, X. Liu, X. Cao et al., Artificial intelligence: a powerful paradigm for scientific research. *The Innovation* **2**, 100179 (2021). <https://doi.org/10.1016/j.xinn.2021.100179>
12. A. Boehnlein, M. Diefenthaler, N. Sato et al., Colloquium: machine learning in nuclear physics. *Rev. Mod. Phys.* **94**, 031003 (2022). <https://doi.org/10.1103/RevModPhys.94.031003>
13. V. Singh, S. Patra, N.A. Murugan et al., Recent trends in computational tools and data-driven modeling for advanced materials. *Adv. Mater.* **3**, 4069–4087 (2022). <https://doi.org/10.1039/D2MA00067A>
14. L. Guo, D. Zhao, G. Du et al., Fluorescence turn-on mode of Eu3+ complex nanocomposite to detect histamine for seafood freshness. *Spectrochim. Acta. A Mol. Biomol. Spectrosc.* **302**, 123089 (2023). <https://doi.org/10.1016/j.saa.2023.123089>
15. M.S. El Tokhy, Rapid and robust radioisotopes identification algorithms of X-Ray and gamma spectra. *Measurement* **168**, 108456 (2021). <https://doi.org/10.1016/j.measurement.2020.108456>
16. B. Jeon, S. Lim, E. Lee et al., Deep learning-based pulse height estimation for separation of pile-up pulses from NaI (TI) detector.

- IEEE Trans. Nucl. Sci. **69**, 1344–1351 (2021). <https://doi.org/10.1109/TNS.2021.3140050>
17. A. Akhavanallaf, I. Shiri, H. Arabi et al., Whole-body voxel-based internal dosimetry using deep learning. *Eur. J. Nucl. Med. Mol. Imaging* **48**, 670–682 (2021). <https://doi.org/10.1007/s00259-020-05013-4>
  18. J. Griffiths, S. Kleinegesse, D. Saunders et al., Pulse shape discrimination and exploration of scintillation signals using convolutional neural networks. *Mach. Learn. Sci. Technol.* **1**, 045022 (2020). <https://doi.org/10.1088/2632-2153/abb781>
  19. B. Jeon, J. Kim, E. Lee et al., Pseudo-gamma spectroscopy based on plastic scintillation detectors using multitask learning. *Sensors* **21**, 684 (2021). <https://doi.org/10.3390/s21030684>
  20. A. Regadio, L. Esteban, S. Sanchez-Prieto, Unfolding using deep learning and its application on pulse height analysis and pile-up management. *Nucl. Instrum. Meth. A* **1005**, 165403 (2021). <https://doi.org/10.1016/j.nima.2021.165403>
  21. S. Woldegiorgis, A. Enqvist, J. Baciak, ResNet and CycleGAN for pulse shape discrimination of He-4 detector pulses: recovering pulses conventional algorithms fail to label unanimously. *Appl. Radiat. Isot.* **176**, 109819 (2021). <https://doi.org/10.1016/j.apradiso.2021.109819>
  22. L. Janjanam, S.K. Saha, R. Kar et al., Volterra filter modelling of non-linear system using artificial electric field algorithm assisted Kalman filter and its experimental evaluation. *ISA Trans.* **125**, 614–630 (2022). <https://doi.org/10.1016/j.isatra.2020.09.010>
  23. V.T. Jordanov, K.V. Jordanova, Unfolding-synthesis technique for digital pulse processing, part 2: synthesis. *Nucl. Instrum. Meth. A* **1044**, 167421 (2022). <https://doi.org/10.1016/j.nima.2022.167421>
  24. H.Q. Zhang, Z.D. Li, B. Tang et al., Optimal parameter choice of CR-RC m digital filter in nuclear pulse processing. *Nucl. Sci. Tech.* **30**, 108 (2019). <https://doi.org/10.1007/s41365-019-0638-7>
  25. W.G. Song, L.J. Zhang, G.Y. Wang, A method to restrain parameter drift in trapezoidal pulse shaping. *IEEE Trans. Nucl. Sci.* **67**, 1710–1714 (2020). <https://doi.org/10.1109/TNS.2020.2995901>
  26. W.G. Song, L.J. Zhang, G.Y. Wang et al., Optimized digital Sallen-Key shaping algorithm for radiation detector signal processing. *Nucl. Technol.* **207**, 292–298 (2021). <https://doi.org/10.1080/00295450.2020.1747838>
  27. M. Wang, J.B. Zhou, X.P. Ouyang et al., Gaussian shaper for nuclear pulses based on multilevel cascade convolution. *Nucl. Sci. Tech.* **33**, 160 (2022). <https://doi.org/10.1007/s41365-022-01145-4>
  28. D. Stoller, S. Ewert, S. Dixon, Wave-u-net: a multi-scale neural network for end-to-end audio source separation. *arXiv preprint arXiv:1806.03185* (2018). <https://doi.org/10.48550/arXiv.1806.03185>
  29. L. Tang, X.K. Ma, K.B. Shi et al., A method for correcting characteristic X-ray net peak count from drifted shadow peak. *Nucl. Sci. Tech.* **34**, 175 (2023). <https://doi.org/10.1007/s41365-023-01333-w>
  30. X.K. Ma, H.Q. Huang, X. Ji et al., X-ray spectra correction based on deep learning CNN-LSTM model. *Measurement* **199**, 111510 (2022). <https://doi.org/10.1016/j.measurement.2022.111510>
  31. C. Gao, P. Zhao, Q. Fan et al., Deep neural network: as the novel pipelines in multiple preprocessing for Raman spectroscopy. *Spectrochim. Acta. A Mol. Biomol. Spectrosc.* **302**, 123086 (2023). <https://doi.org/10.1016/j.saa.2023.123086>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Lin Tang<sup>1,2,3,4</sup>  · Shuang Zhou<sup>1</sup> · Kai-Bo Shi<sup>1</sup> · Hong-Tao Shen<sup>2</sup> · Lei You<sup>5,6</sup>

✉ Kai-Bo Shi  
skbs111@163.com

✉ Hong-Tao Shen  
shenht@gxnu.edu.cn

<sup>1</sup> College of Electronic Information and Electrical Engineering, Chengdu University, Chengdu 610106, China

<sup>2</sup> Guangxi Key Laboratory of Nuclear Physics and Nuclear Technology, Guangxi Normal University, Guilin 541004, China

<sup>3</sup> National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230039, China

<sup>4</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

<sup>5</sup> Vanadium and Titanium Resource Comprehensive Utilization Key Laboratory of Sichuan Province, Panzhihua University, Panzhihua 617000, China

<sup>6</sup> Key Laboratory of Interior Layout optimization and Security, Institutions of Higher Education of Sichuan Province, Chengdu Normal University, Chengdu 611130, China