



# A method for correcting characteristic X-ray net peak count from drifted shadow peak

Lin Tang<sup>1,2,3,4</sup> · Xing-Ke Ma<sup>5</sup> · Kai-Bo Shi<sup>1</sup> · Yeng-Chai Soh<sup>4</sup> · Hong-Tao Shen<sup>2</sup>

Received: 18 May 2023 / Revised: 11 July 2023 / Accepted: 14 July 2023 / Published online: 21 November 2023

© 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 2023

## Abstract

To correct spectral peak drift and obtain more reliable net counts, this study proposes a long short-term memory (LSTM) model fused with a convolutional neural network (CNN) to accurately estimate the relevant parameters of a nuclear pulse signal by learning of samples. A predefined mathematical model was used to train the CNN–LSTM model and generate a dataset composed of distorted pulse sequences. The trained model was validated using simulated pulses. The relative errors in the amplitude estimation of pulse sequences with different degrees of distortion were obtained using triangular shaping, CNN–LSTM, and LSTM models. As a result, for severely distorted pulses, the relative error of the CNN–LSTM model in estimating the pulse parameters was reduced by 14.35% compared with that of the triangular shaping algorithm. For slightly distorted pulses, the relative error of the CNN–LSTM model was reduced by 0.33% compared with that of the triangular shaping algorithm. The model was then evaluated considering two performance indicators, the correction ratio and the efficiency ratio, which represent the proportion of the increase in peak area of the two characteristic peak regions of interest (ROIs) to the peak area of the corrected characteristic peak ROI and the proportion of the increase in peak area of the two characteristic peak ROIs to the peak areas of the two shadow peak ROI, respectively. Ten measurement results of the iron ore samples indicate that approximately 86.27% of the decreased peak area of the shadow peak ROI was 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 1.72%. The proposed CNN–LSTM model can be applied to X-ray energy spectrum correction, which is of great significance for X-ray spectroscopy and elemental content analyses.

**Keywords** Peak correction · Triangular shaping · Deep learning · Long short-term memory · Convolutional neural network · X-ray fluorescence spectroscopy · Silicon drift detector

## 1 Introduction

X-ray fluorescence refers to the X-rays emitted by a sample under irradiation by an excitation source, which contains the elemental and chemical composition information of the analyzed sample. In X-ray fluorescence (XRF) spectrometry, the counting rate and energy resolution are important indicators that directly determine the accuracy of the content analysis of each element in the tested sample [1]. In particular, in the detection of weak elements with lower contents, peak drift and count loss have an inestimable impact on the measurement results. The main causes of peak drift and count loss are pulse distortions caused by the measurement system itself. The key elements of the measurement system include a probe (integrated with the detector and preamplifier), X-ray tube, tested sample, front-end signal conditioning circuit,

---

This work was supported by the Open Project of the Guangxi Key Laboratory of Nuclear Physics and Nuclear Technology (No. NLK2022-05), Central Government Guidance Funds for Local Scientific and Technological Development, China (No. Guike ZY22096024), Sichuan Natural Science Youth Fund Project (No. 2023NSFSC1366), Open Research Fund of the National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University (No. AE202209), Research Fund of Guangxi Key Lab of Multi-source Information Mining & Security (MIMS22-04), and National Natural Science Youth Foundation of China (No. 12305214).

---

Extended author information available on the last page of the article

digital processing unit, controller unit, and upper computer [2]. Distorted pulses primarily include stacked, interfering, slow, spark, double, and truncated pulses. In measurement systems using switch reset preamplifiers, distorted pulses are mainly composed of truncated pulses, which refer to a pulse signal whose pulse amplitude suddenly jumps to zero owing to the reset of the switch, resulting in an insufficient effective width. As a result, the amplitude loss of the triangular shaping results caused by pulse distortion has led to some limitations in current X-ray fluorescence spectroscopy, including spectral peak drift, unreliable net counts, and inaccurate element content analysis.

In the field of X-ray fluorescence spectroscopy, research has mainly focused on digital pulse shaping [3] and filtering [4], and an increasing number of researchers are focusing on using new digital signal processing technologies to solve problems in this field. The algorithm proposed by Zhong [5] solved the problem of poor resolution of the energy spectrum caused by pulse stacking and temperature fluctuations in an X-ray spectrum system. A symmetric conversion method based on Gaussian distribution [6] was proposed to obtain the  $\gamma$ -ray net count from the interlaced overlap peak in the HPGe  $\gamma$ -ray spectrometer system. A modified sparse reconstruction method [7] to overcome pulse pile-up, especially with ultrahigh count rates, which uses two regularization terms to compensate for the error caused by an inadequate sampling rate. To achieve high count rates, a new true Gaussian digital shaper for detector pulses [8] and a compensation technology for pulse stacking [9] were proposed. In our previous research, we proposed a pulse elimination method [10] and pulse repair method [11] for distorted pulses, both of which improved the accuracy of spectral analysis to a certain extent. As traditional pulse processing methods, the above research methods have obvious optimization effects on the X-ray fluorescence spectrum analysis in situations where pulse stacking or pulse distortion is not particularly serious. However, traditional pulse processing methods are significantly limited when pulse stacking or pulse distortion is difficult to recognize, and this has become a popular research topic in the field of spectrum processing.

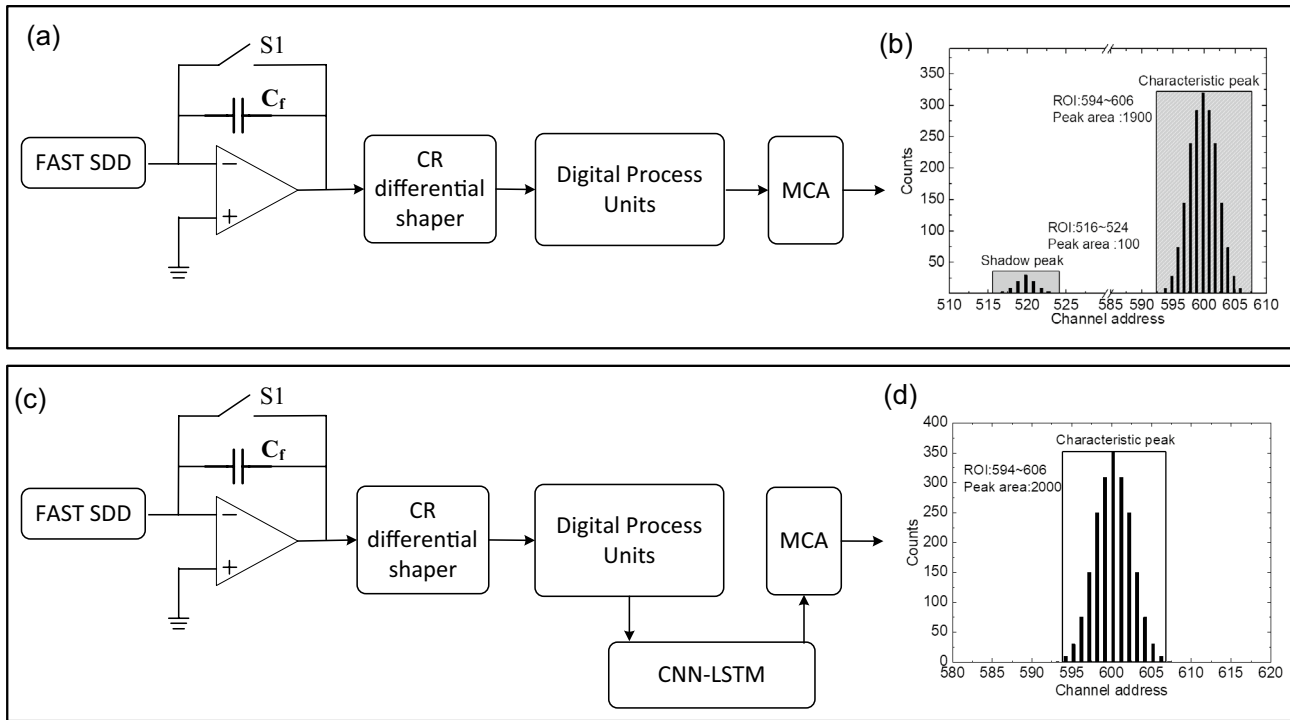
In recent years, deep learning technology has developed rapidly, and many excellent models have emerged, such as UNet [12], VNet [13], and Transformer [14], which have been widely used in medicine [15], industry [16], control [17], and radiation measurements, such as imaging quality improvement in radiation therapy [18], gamma spectrum analysis [19], and pulse signal analysis based on residual structures [20]. Deep learning technology provides various ideas for pulse processing in radiation measurement [21, 22]. Touch [23] applied an artificial neural network for energy spectrum correction and achieved satisfactory results. Alberto et al. [24] proposed a specific type of UNet that

filters pulses, returns their height, and estimates the pulse amplitude. Byoungil et al. [25] proposed a deep learning-based method for separating and predicting the true pulse height of a signal for application in spectroscopy with a scintillation detector. Liu [26] investigated a pulse-coupled neural network (PCNN) for higher anti-noise performance in the neutron and gamma-ray ( $n-\gamma$ ) discrimination field. Ma [27, 28] accurately predicted the trapezoidal-forming parameters of stacked pulses using a long short-term memory (LSTM) model. Based on previous research on pulse amplitude estimation, this paper proposes a methodology for an LSTM model fused with a convolutional neural network (CNN). Compared to other pulse estimation algorithms, this algorithm exhibited better performance. The introduction and extension of this new technology to X-ray fluorescence spectroscopy are of significant interest.

## 2 Principle and method

### 2.1 Principle of peak drift

The X-ray fluorescence spectrum is often analyzed using multichannel pulse amplitude (MCA), with each pulse amplitude corresponding to a count in the counting histogram. When the pulse output of the measurement system experienced an amplitude loss during the digital processing stage, the corresponding counting histogram of the pulse drifted to the left. When the number of pulses is sufficient, they exist as shadow peaks of the characteristic peaks in the generated energy spectrum. The traditional spectrum acquisition process is shown in Fig. 1a and includes a detector, a preamplifier, a CR differential shaper, a digital processing unit, and an MCA unit. For standard sources with a single-element composition, the number of characteristic peaks is limited; therefore, the pulse amplitude of the measured output mostly fluctuates within a certain range. Taking 2000 pulses with an amplitude of approximately 600 mV as an example and assuming a pulse distortion ratio of 5%, the X-ray spectrum obtained using the spectral analysis method shown in Fig. 1a is shown in Fig. 1b. In this figure, the channel range of the characteristic peak region of interest (ROI) is 594–606 with a peak area (net peak count) of 1900. A shadow peak formed by the distorted pulses appeared near the 520th channel on the left side of the characteristic peak, with a peak area of 100, as shown by the green shaded area in Fig. 1b. In practical applications, if the counting rate of the characteristic peak is high, the number of distorted pulses accumulates to a certain extent and a shadow peak is generated. This shadow peak not only reduces the net count of the characteristic peak ROI but also introduces new difficulties to spectral analysis. Therefore, this study proposes a deep learning-based CNN–LSTM model that is added before the MCA unit to achieve an accurate estimation of pulse parameters.



**Fig. 1** Principle of peak drift and correction. **a** Spectrum acquisition process of the traditional method, **b** the X-ray spectrum by the traditional method, **c** spectral acquisition process of the CNN–LSTM model, **d** the histogram of the characteristic peaks by calling the CNN–LSTM model

The spectral acquisition process for the added model is shown in Fig. 1c. In an ideal situation, when the amplitudes of 100 distorted pulses are accurately estimated, the histogram of the characteristic peaks obtained by calling the model is shown in Fig. 1d, where the ROI of the characteristic peaks remains unchanged but the total count increases to 2000. It can be concluded that calling the model to optimize the pulse amplitude estimation not only ensures that the counting of characteristic peaks is not lost, but also eliminates shadow peaks caused by pulse distortion.

## 2.2 Deep learning

### 2.2.1 Data acquisition

In the data acquisition stage, the datasets were produced. For a single distorted negative exponential pulse, the pulse distortion time is assumed to be  $t_{\text{jump}}$ , and its mathematical model is given by Eq. (1), where  $A$  represents the amplitude of the negative exponential pulse,  $\tau$  represents attenuation time constant, and  $T_{\text{clk}}$  represents sampling period:

$$V_e(t) = \begin{cases} A \times \exp\left(-t \times \frac{T_{\text{clk}}}{\tau}\right), & t < t_{\text{jump}} \\ 0, & t \geq t_{\text{jump}} \end{cases} \quad (1)$$

The mathematical model of a pulse sequence composed of  $N$  distortion-negative exponential pulses is:

$$V_e(t) = \sum_{i=1}^N \left[ u(t - T_i) A_i e^{-\frac{(t-T_i)}{\tau}} \right]. \quad (2)$$

Here,  $u(t)$  represents the unit-step signal,  $A_i$  is the amplitude coefficient of the  $i$ th nuclear pulse,  $T_i$  is the occurrence time of the  $i$ th nuclear pulse, and  $\tau$  represents the time constant. The negative exponential pulse sequence after discretization can be expressed as

$$V_e(kT_{\text{clk}}) = \sum_{i=1}^N \left[ u(kT_{\text{clk}} - T_i) A_i e^{-\frac{(kT_{\text{clk}}-T_i)}{\tau}} \right]. \quad (3)$$

The distorted nuclear pulse sequence  $V_o(nT_{\text{clk}})$  used for the parameter estimation is regarded as  $N$  distorted negative exponential pulse sequences  $V_e(kT_{\text{clk}})$  obtained after triangular shaping, and its mathematical model is as follows:

$$V_o(nT_{\text{clk}}) = 2V_o[(n-1)T_{\text{clk}}] - V_o[(n-2)T_{\text{clk}}] + \frac{1}{n_a} \{ V_e[(n-1)T_{\text{clk}}] - V_e[(n-n_a-1)T_{\text{clk}}] \} - e^{-\frac{T_{\text{clk}}}{\tau}} \{ V_e[(n-2)T_{\text{clk}}] - V_e[(n-n_a-2)T_{\text{clk}}] \}. \quad (4)$$

Here, the dataset of the CNN–LSTM model established in this study was taken from the pulse amplitude sampling value of distortion pulses after triangular shaping, whereas the parameter set  $P$  was taken from the negative exponential

pulses before triangular shaping. Taking the parameter set  $P_i$  of the  $i$ th negative exponential pulse as an example, this set includes the amplitude  $A_i (i = 1, 2, \dots, N)$ , sampling period  $T_{clk}$ , time constant  $\tau$  of the negative exponential pulse, and the rising time  $t_{up}$  of the triangular shaping. The matrix representation of the dataset is as follows:

$$\begin{bmatrix} [V_o(T_{clk})]_1 & [V_o(2 \times T_{clk})]_1 & \dots & [V_o(n \times T_{clk})]_1 P_1 \\ \vdots & \vdots & \ddots & \vdots \\ [V_o(T_{clk})]_N & [V_o(2 \times T_{clk})]_N & \dots & [V_o(n \times T_{clk})]_N P_N \end{bmatrix} \quad (5)$$

The datasets in Eq. (5) contain  $N$  triangular-shaped pulse sequences, and the amplitude of each pulse corresponds to a row in the matrix. Each row contains  $n + 1$  columns. The first  $n$  columns correspond to each amplitude value of the triangular shaping result of distortion pulses, and the last column represents the parameter set of the sequence, including  $A_i, T_{clk}, \tau$ , and  $t_{up}$ . The structures of the generated datasets are shown in Fig. 2.

As required, this study divided the dataset into training, test, and validation sets in a ratio of 7:2:1. In general, the training set accounts for a large proportion of the data and is used to train the generalization ability of the model, whereas the verification set is used to verify whether the model is overfitted. Once overfitting occurs, it must be eliminated by adding a dropout layer to randomly discard the connections of some neurons.

### 2.2.2 Hyperparameter optimization

The model adopted in this study addresses a series of nuclear pulse amplitude sequences. As a special recurrent neural network (RNN), LSTM is essentially different from an RNN in that the introduction of forgetting gates determines which information will be retained or forgotten by controlling the parameters. Therefore, LSTM can solve the problems of gradient disappearance and explosion of long time-series samples during the training process.

LSTM usually deals with long sequences and large sample data, but too large sample data will cause some difficulties during model training, such as computation complexity. Therefore, this study combined LSTM with a CNN. A convolutional neural network (CNN) is not as much of an algorithm as a feature extraction method and usually includes a convolution layer (including an activation layer), a pooling layer, and a full link layer. The process of extracting features through convolutional neural network is essentially the process of solving the optimal parameter matrix. The relationship between input and output can be expressed using Eq. (6):

$$y = f(X, W) = WX + b, \quad (6)$$

where  $W$  represents the weight parameter matrix,  $X$  represents the input neuron matrix, as shown in Eq. (6), and  $b$  represents the offset term. The structures and parameter settings of the CNN model are listed in Table 1. By setting multiple convolution and pooling layers, the CNN greatly reduced the number of samples in the dataset while completing feature

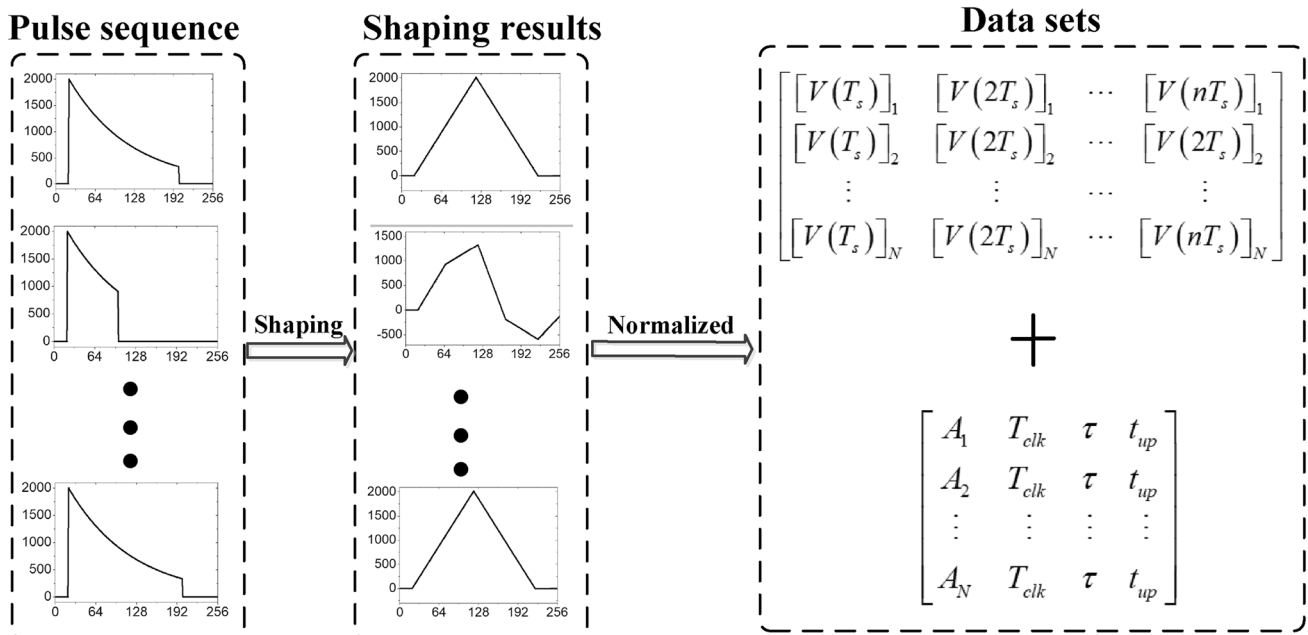


Fig. 2 Generated datasets

**Table 1** CNN architecture details

Block	Layer (filter size)	Input size	Output size
Conv1D_1	Conv1D (3×3)	(None, 1, 256)	(None, 1, 64)
Max_pooling1D_1	Max_pooling1D	(None, 1, 64)	(None, 1, 64)
Conv1D_2	Conv1D (3×3)	(None, 1, 64)	(None, 1, 16)
Max_pooling1D_2	Max_pooling1D	(None, 1, 16)	(None, 1, 16)

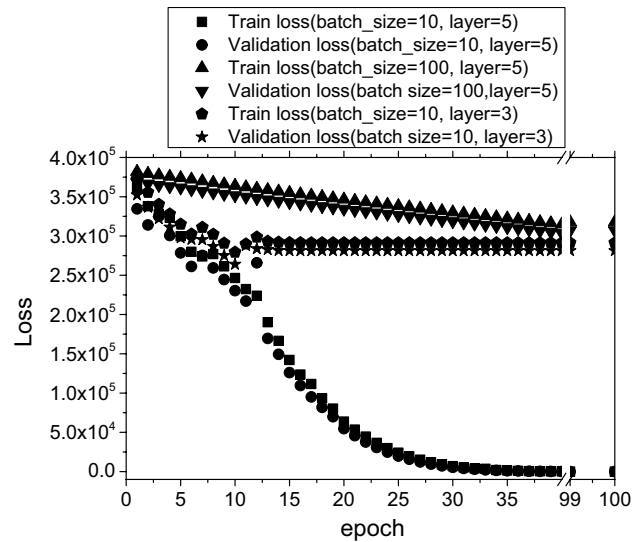
extraction. The CNN model used in this study included two one-dimensional convolution layers and two pooling layers. The time step for each layer was set to one. The first convolution layer contained 64 convolution cores and output 64 eigenvectors, whereas the second contained 16 convolution kernels and output 16 eigenvectors. The size of the convolution core in both convolution layers is set to 3×3, the moving step is 1, the filling strategy is "same," and the activation function is "relu." The pooling layer adopts MaxPooling1D, which does not change the input signal size.

In the forward propagation process, the input neurons are kept unchanged, and the weight parameter matrix is initialized using a random strategy. The error between the output parameter set  $P'_i$  in the forward propagation process and the actual pulse parameter set  $P_i$  in the training set can be calculated using the loss function. The calculation method is expressed in Eq. (7), where the actual parameter set of the  $i$ th sample is represented by  $P_i$  and  $P_i$  the parameter set estimated by the forward propagation is represented by  $P'_i$ . For the training set with  $N$  samples, the mean square error (MSE) of the parameter set was considered as the function value of the loss function, which is represented by  $L_{MSE}$ .

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (P_i - P'_i)^2 \tag{7}$$

Subsequently, the backpropagation through time (BPTT) algorithm is applied to feed back the gradient of the loss function and  $L_{MSE}$  to the CNN–LSTM network to update the weight matrix  $W$ , so as to reduce the error in subsequent iterations. To prevent a gradient explosion, this model sets the gradient clipping parameter clipnorm to 1 and clipvalue to 0.5.

Figure 3 shows the hyperparameter optimization process for the parameters and layers during the training process of the CNN–LSTM model. If the batch size is set too large, it can lead to memory overflow during the training process, and the model is prone to convergence to local optima, making it impossible to complete the training. If the batch size is too small, the rate of convergence of the model will be too slow, and the training time will be too long. Figure 3 shows the iterative loss values obtained for the training and validation sets when the number of layers of the LSTM model was five and the parameter batch size was set to 10 and 100, respectively. When the batch size was 100, the model



**Fig. 3** Process of the hyperparameter optimization

converged at the 40th epoch with a high loss of  $3 \times 10^5$ . When batch size is 10, the model converges normally, and the loss value after convergence approaches zero as much as possible.

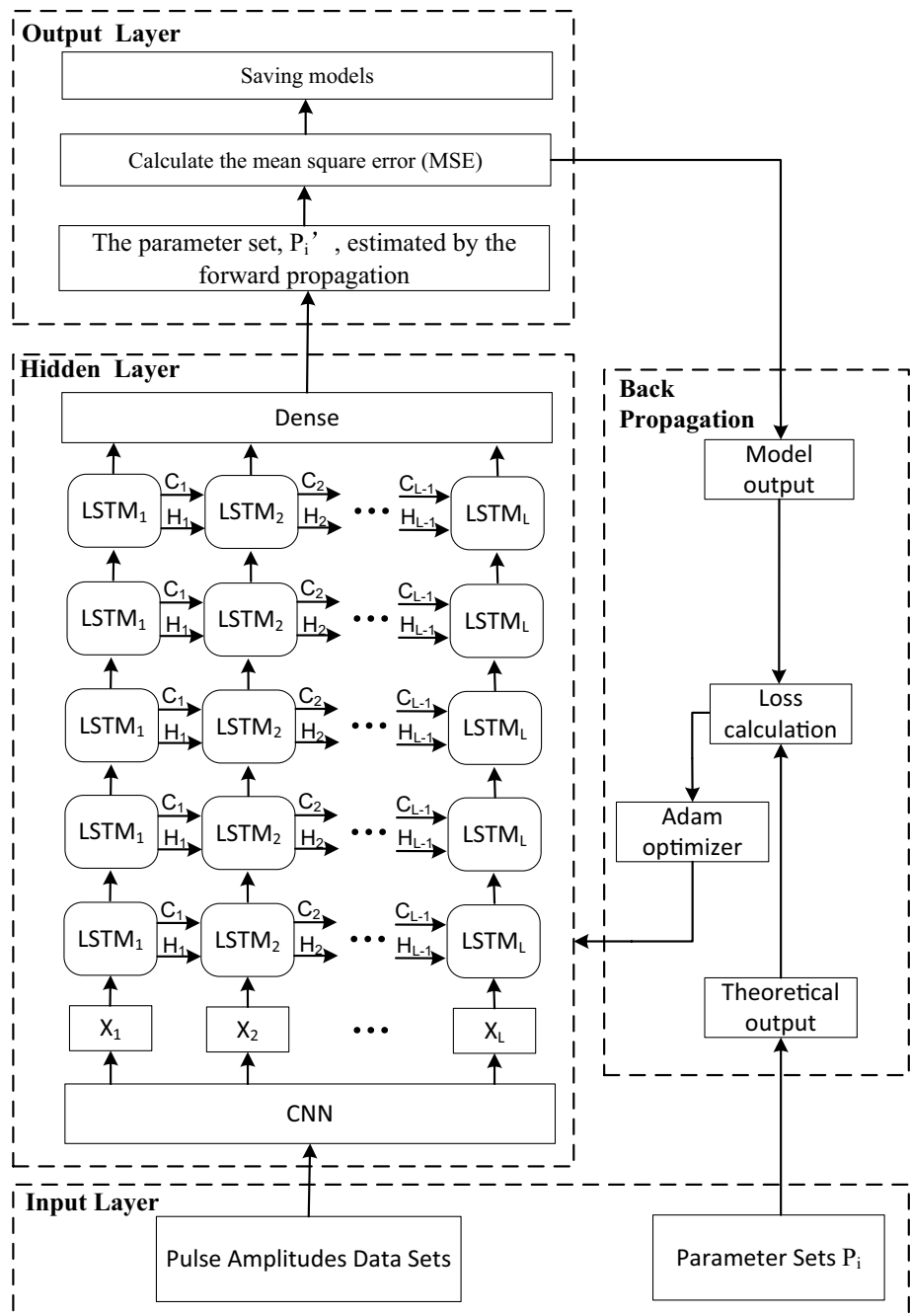
When setting the parameters for the LSTM model, theoretically speaking, the more layers there are, the more ideal the training results. The problem of vanishing gradient must also be considered. An increase in the number of layers results in a greater computational burden; therefore, when optimizing the hyperparameters, we usually set the number of layers to 3–6. Figure 3 shows the iterative loss values obtained for the training and validation sets for layers 3 and 5. It can be seen that when the batch size is 10, the attenuation speed of the LSTM model with five and three layers is close, but when the layer number is 3, the loss value after model convergence is still as high as  $2.8 \times 10^5$ . When the number of layers was five, the loss values of the training and validation sets approached zero.

After hyperparameter optimization, five LSTM layers were set with an initial learning rate of 0.0001 and a batch size of 10, and Adam was selected as the optimizer. The generated network model structure is shown in Fig. 4, which includes the input layer, hidden layer, output layer, and back-propagation part. The hidden layer includes the CNN and LSTM models.

### 3 Simulation results and experimental verification

The CNN–LSTM model proposed in this study was applied to the peak correction of the X-ray fluorescence spectrum. As mentioned above, when the negative exponential pulse

**Fig. 4** Network model structure of CNN–LSTM. The detailed calculation procedure of the LSTM unit can be found in the article of Graves et al. [29]



sequence output of the measurement system was significantly distorted, the amplitude value of the triangular shaping result was significantly damaged. According to the generation principle of the digital multichannel spectrum, the amplitude loss of the distorted pulses after shaping appears in the form of count drift in the X-ray fluorescence spectrum, which is unfavorable for obtaining an accurate X-ray fluorescence spectrum. The CNN–LSTM model proposed in this study, based on deep learning, aims to accurately estimate the parameters of the triangular shaping results of the distorted pulses. Thus, the shift in the peak in the X-ray

fluorescence spectrum can be corrected to obtain a more accurate X-ray fluorescence spectrum.

### 3.1 CNN–LSTM simulation

#### 3.1.1 Model training

To verify the effect of the CNN–LSTM model on the parameter estimation of the triangular shaping results under the condition of severe distortion of the negative exponential pulses, we took 10,000 samples and divided

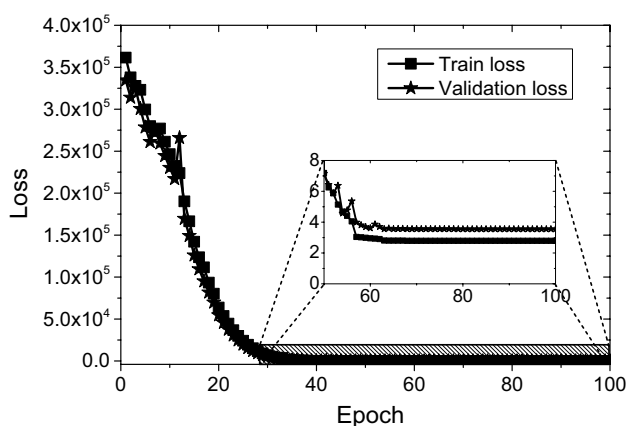
them into training, verification, and test sets according to a ratio of 7:2:1, with a training period of 100 epochs. The change in the loss values obtained from the training and verification sets during the training process is shown in Fig. 5.

In general, with an increase in the number of training cycles, the loss values of the training and verification sets showed a downward trend. The verification set experiences some shocks in the later period, but soon tends to stabilize. Using the loss function, when the loss values of both the training and validation sets were low and relatively stable, the model during that period was saved as the best model. In the training process of the proposed model, the model at the 91st epoch was saved as the best model, with loss values of 2.7895 and 3.5507 for the training and validation sets, respectively, during this epoch.

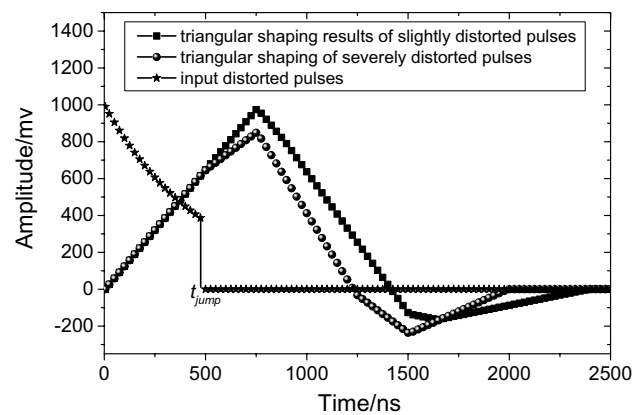
### 3.1.2 Performance evaluation of parameter estimation

In the production of the test set, considering that the distortion degree of the pulses may affect the model test effect, the sample of the test set can be divided into two categories according to the pulse distortion time mentioned above: slightly distorted pulses and severely distorted pulses, whose triangular shaping results are shown in Fig. 6. The rising time of the triangular shape,  $t_{up}$ , is bound. If the time of pulse distortion,  $t_{jump}$ , is before the  $t_{up}$ , the amplitude of the shaping result decreases significantly, as shown in Fig. 6, and then such pulses are marked as severely distorted pulses. On the other hand, if the time of pulse distortion is after  $t_{up}$  (including the critical time  $t_{up}$ ), the amplitude of the shaping result decreases slightly, as shown in Fig. 6; therefore, this type of pulse is marked as a slightly distorted pulse.

To control the impact of other variables, the amplitude parameter  $A_i$  of Pulse1–Pulse8 was fixed at 2000 mV, the time interval  $T_i$  of adjacent distorted pulses was  $500 T_{clk}$ , the



**Fig. 5** Iterative graph of loss and accuracy on training and validation sets during model training



**Fig. 6** Triangular shaping results of negative exponential pulse at different distortion moments

sampling period  $T_{clk}$  was 50 ns, and the rising time of the triangular shaping was  $100 T_{clk}$ . Consider two pulse sequences and their triangular shaping results for analysis, as shown in Figs. 7 and 8.

In Fig. 7a, the amplitude value of Pulse1 suddenly changes to 0 at the 47th  $T_{clk}$  and the amplitude value of Pulse2 suddenly changes to 0 at the 60th  $T_{clk}$ . The amplitude value of Pulse3 suddenly changes to zero at the 75th  $T_{clk}$ , and the amplitude value of Pulse4 suddenly changes to zero at the 90th  $T_{clk}$ . As mentioned above, the rising time of the triangular shaping remains at  $100 T_{clk}$ , that is, Pulse1–Pulse 4 are distorted before the peak value of the triangular shaping. Therefore, the amplitude values of the shaping results exhibited a significant loss compared to the original pulses, as shown in Fig. 7b.

In Fig. 8a, the amplitude value of Pulse5 suddenly changes to zero at the 100th  $T_{clk}$  and the amplitude value of Pulse6 suddenly changes to zero at the 140th  $T_{clk}$ . The amplitude value of Pulse7 suddenly changes to zero at the 220th  $T_{clk}$  and the amplitude value of Pulse8 suddenly changes to zero at the 280th  $T_{clk}$ . In Fig. 8a, the common feature of Pulse5–Pulse8 is that the distortion time is after the peak value of the triangular shaping. Therefore, the amplitude values of the shaping results have no large losses compared to the original pulses, as shown in Fig. 8b.

During the performance evaluation of the parameter estimation, the CNN–LSTM and LSTM models were used to estimate the parameters of the above two pulse sequences, and the output results are shown in Fig. 9. The real value of the eight pulse amplitudes was fixed at 2000 mV, but the amplitude loss of the triangular shaping result of the Pulse 1–Pulse 4 was very large.

Based on the parameter estimation results of the slightly and severely distorted pulses, the absolute and relative errors of the different methods used to estimate the pulse amplitudes are summarized in Table 2.

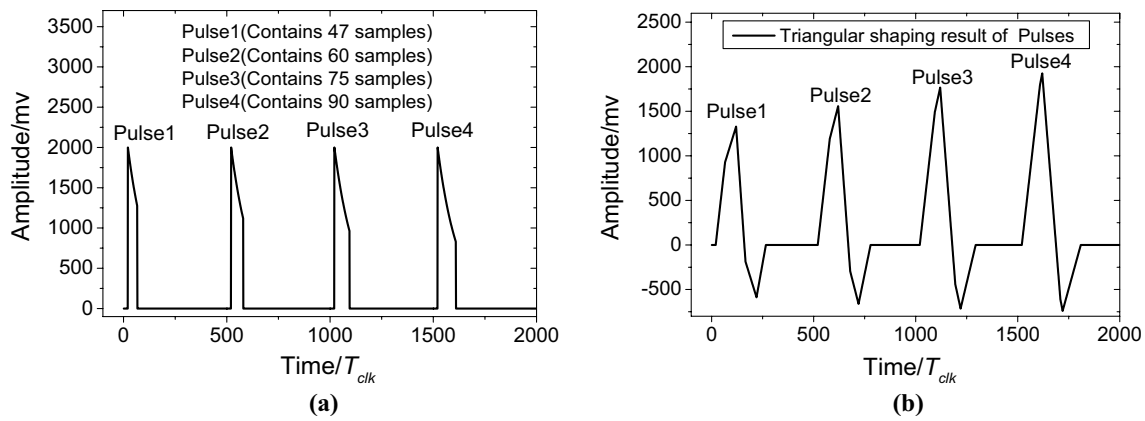


Fig. 7 Triangular shaping results of the seriously distorted pulses. **a** Seriously distorted pulses, **b** triangular shaping results

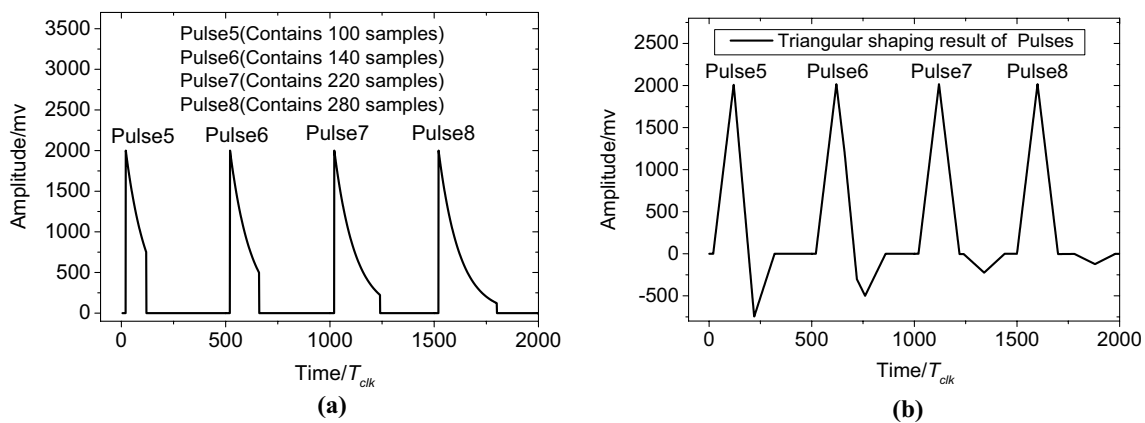


Fig. 8 Triangular shaping results of the slightly distorted pulses. **a** Slightly distorted pulses, **b** triangular shaping results

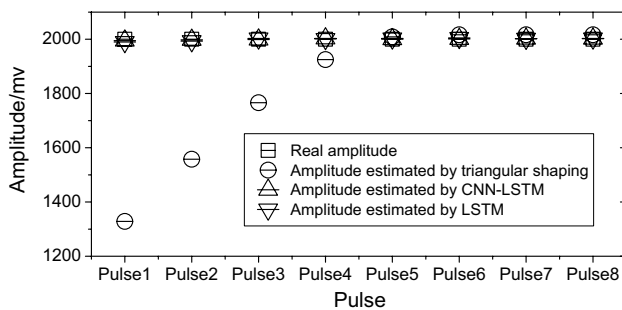


Fig. 9 Amplitude comparison chart of the distorted pulses

$A_{real}$  represents the actual pulse amplitude. The measurement methods used in this study mainly included triangular shaping, CNN-LSTM, and LSTM models, the measurement results of which are represented by  $A_{Tri}$ ,  $A_{CNN-LSTM}$ , and  $A_{LSTM}$ . This study used relative error indicators to evaluate the parameter estimation performance of each algorithm. Let  $\delta$  represent the relative error

and  $\Delta$  represent the absolute error, whose calculation process is given in Eqs. (8), (9), and (10).

$$\delta_{Tri} = \frac{\Delta_{Tri}}{A_{real}} \times 100\% = \frac{ABS(A_{real} - A_{Tri})}{A_{real}} \times 100\% \tag{8}$$

$$\begin{aligned} \delta_{CNN-LSTM} &= \frac{\Delta_{CNN-LSTM}}{A_{real}} \times 100\% \\ &= \frac{ABS(A_{real} - A_{CNN-LSTM})}{A_{real}} \times 100\% \end{aligned} \tag{9}$$

$$\delta_{LSTM} = \frac{\Delta_{LSTM}}{A_{real}} \times 100\% = \frac{ABS(A_{real} - A_{LSTM})}{A_{real}} \times 100\% \tag{10}$$

For severely distorted pulses, the average relative error of triangular shaping was as high as 14.39%, whereas that of CNN-LSTM was only 0.14%, and that of LSTM was 1.05%. On the other hand, for slightly distorted pulses, the average relative error of triangular shaping was 0.39%, while



**Table 2** Comparison of estimated values obtained by different models

	Seriously distorted pulses				Slightly distorted pulses			
	Pulse1	Pulse2	Pulse3	Pulse4	Pulse5	Pulse6	Pulse7	Pulse8
$A_{real}(mV)$	2000	2000	2000	2000	2000	2000	2000	2000
$A_{Tri}(mV)$	1459.7	1629	1820	1940	1998.4	2016.4	2008.6	2004.4
$\Delta_{Tri}$	540.3	371	180	60	1.6	16.4	8.6	4.4
$\delta_{Tri}$	27.02%	18.55%	9.00%	3.00%	0.08%	0.82%	0.43%	0.22%
$A_{CNN-LSTM}(mV)$	1995	1998	1999	2003	2000	2001	2002	2002
$\Delta_{CNN-LSTM}$	5	2	1	3	0	1	2	2
$\delta_{CNN-LSTM}$	0.25%	0.10%	0.05%	0.15%	0.00%	0.05%	0.10%	0.10%
$A_{LSTM}(mV)$	1990	1993	2003	2001	2003	2005	2002	2003
$\Delta_{LSTM}$	10	7	3	1	3	5	2	3
$\delta_{LSTM}$	0.50%	0.35%	0.15%	0.05%	0.15%	0.25%	0.10%	0.15%

that of CNN–LSTM was only 0.06% and that of LSTM was 0.65%. It can also be observed from the estimation results that the two models can estimate the pulse amplitude very accurately, whether it is a severely or slightly distorted pulse. It is worth noting that although the performance of the CNN–LSTM model is slightly better than that of LSTM, both deep learning methods have very high accuracy in pulse parameter estimation. This study introduces a CNN because the input pulse sequence is relatively complex, and directly uses the LSTM model to process data that are too large. Therefore, using a CNN for sampling reduces the amount of data and saves computational resources.

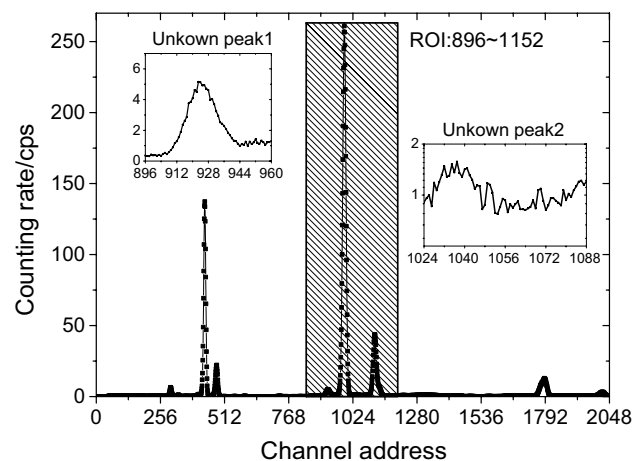
### 3.2 Experimental verification

In the model simulation, we used two types of distortion pulses to train, verify, and test the CNN–LSTM model and achieved good test results in the parameter estimation of distortion-negative exponential pulses. To further verify the optimization effect of the parameter estimation on the X-ray fluorescence spectrum, an iron ore sample was selected as the measurement object in the experimental verification link. In previous studies [11], we identified that the elemental components with high content in this sample were Fe, Sr, and Sn. The measurement system included a high-performance silicon drift detector (FAST-SDD) and KYW2000A X-ray tube. The effective detection area of the detector is 25 mm<sup>2</sup>, the detector thickness is 500 μm, and the thickness of beryllium window is 12.5 μm. The rated tube voltage was 50 kV and the rated tube current was 0–1 mA. The ADC sampling frequency was 20 MHz and the sampling period was 50 ns.

The original measured spectra are shown in Fig. 10. It is easy to see that the element with the highest content in the measured full spectrum is Sr element. According to the principle of distortion-pulse generation, elements with higher counting rates have a higher probability of generating shadowed peaks. Therefore, we selected Sr with

the highest content and considered the net count of the two characteristic peaks of that element and their corresponding shadow peaks in the ROI as the analysis object. There is an unknown peak on the left side of the Sr characteristic peaks. If such shadow peaks are not processed, they may be mistaken as the characteristic peaks of some elements. Unreliable net counts in the ROI directly affect the elemental content analysis. Therefore, it is necessary to correct the shadow peaks in X-ray fluorescence spectra.

In this study, the optimization of the X-ray fluorescence spectrum was implemented using the CNN–LSTM model based on deep learning to correct the count drift caused by distorted pulses. The negative exponential pulse sequence from the measurement process and its triangular shaping results were saved to create the test set. To create the test set, it is necessary to preprocess the negative exponential pulse sequence. To match the trained CNN–LSTM model, the preprocessing unit primarily completes the discrimination and separation of distorted pulses. For experimental verification, qualitative analysis of the X-ray fluorescence



**Fig. 10** Measurement spectrum of iron ore samples

spectrum optimization and quantitative analysis of the spectrum peak correction were completed.

### 3.2.1 Qualitative analysis

The amplitude of the distorted pulses estimated by the CNN–LSTM model was used to replace the original pulse amplitude, and a comparison diagram of the X-ray fluorescence spectrum is shown in Fig. 11. The red spectral lines represent the results after the peak correction. By magnifying the local characteristics of the shadow peak area in the logarithmic coordinate system, it can be observed that there is a weak peak to the left of the two characteristic peaks of strontium in Fig. 11. Because the chemical symbol of Sr is Sr, its two characteristic peaks are represented by Sr-1 and Sr-2. According to the principle of multichannel spectroscopy, the amplitude loss of triangular shaping results in a left shift in the counts (also known as the left shift of the peak position), and the left-shifted counts form a new shadow peak on the left side of the characteristic peak. In Fig. 11, the left shift of the two characteristic peaks of strontium forms the shadow peaks 1 and 2. After the parameter estimation of the distorted pulses using the CNN–LSTM model, the left shift of the peak position was effectively corrected, and the shadow peak was eliminated.

Qualitative analysis of the X-ray fluorescence spectrum before and after optimization showed that the CNN–LSTM model trained in this study could effectively correct the shadow peak caused by pulse distortion and optimize the X-ray fluorescence spectrum analysis results.

### 3.2.2 Quantitative analysis

As mentioned previously, the amplitude loss of shaping results in a left shift in the peak position. A shadow of

the characteristic peak was formed on the left side of the characteristic peak. Here, the peak area  $S$  represents the sum of the counts of a certain channel address interval.  $S_{\text{shadow}}$  represents the area of the shadow peak formed by the left shift of the first characteristic peak of Sr, and  $S_{\text{peakloss}}$  indicates the area loss of the first characteristic peak of strontium owing to the left shift of the peak position, whose computational processes are shown in Eqs. (11), (12), (13), and (14).

$$S_{\text{shadow1}} = \sum_{i=896}^{960} \text{Count}_{i\text{-origin}} - \sum_{i=896}^{960} \text{Count}_{i\text{-corrected}} \quad (11)$$

$$S_{\text{peakloss1}} = \sum_{i=960}^{1024} \text{Count}_{i\text{-corrected}} - \sum_{i=960}^{1024} \text{Count}_{i\text{-origin}} \quad (12)$$

The region of interest (ROI) of shadow peak 1 is 896–960, so  $S_{\text{shadow1}}$  is numerically equal to the difference between the peak area of the channel address interval where the shadow peak is located before and after spectral peak correction, as shown in Eq. (11), and  $S_{\text{shadow1}}$  is shown in the shaded area of Fig. 12a.

$S_{\text{peakloss1}}$  represents the corrected peak area loss in the ROI of the first characteristic Sr peak of strontium element after calling the model. The ROI of the selected characteristic peak was located in the channel address interval of 960–1024, as shown in the shaded area of Fig. 12b.  $S_{\text{peakloss1}}$  is numerically equal to the difference in the peak area of the second characteristic peak ROI before and after the spectral peak correction, as shown in Eq. (12).

$S_{\text{shadow2}}$  represents the area of the shadow peak formed by the left shift of the second characteristic peak of strontium. The ROI of shadow peak 2 is 1024–1088, so  $S_{\text{shadow2}}$  is numerically equal to the difference in the peak area of the shadow peak ROI before and after spectral peak correction, as shown in Eq. (13), and  $S_{\text{shadow2}}$  is shown in the shaded area in Fig. 12c.

$$S_{\text{shadow2}} = \sum_{i=1024}^{1088} \text{Count}_{i\text{-origin}} - \sum_{i=1024}^{1088} \text{Count}_{i\text{-corrected}} \quad (13)$$

$$S_{\text{peakloss2}} = \sum_{i=1088}^{1152} \text{Count}_{i\text{-corrected}} - \sum_{i=1088}^{1152} \text{Count}_{i\text{-origin}} \quad (14)$$

$S_{\text{peakloss2}}$  represents the corrected peak area loss of the second characteristic peak of Sr element after calling the model. The ROI of the selected characteristic peak was located in the channel address interval of 1088–1152, as shown by the shaded area in Fig. 12d.  $S_{\text{peakloss2}}$  is numerically equal to the difference in the peak area of the second

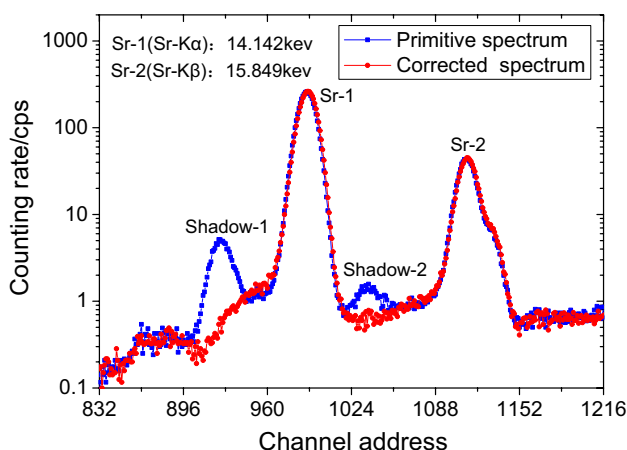
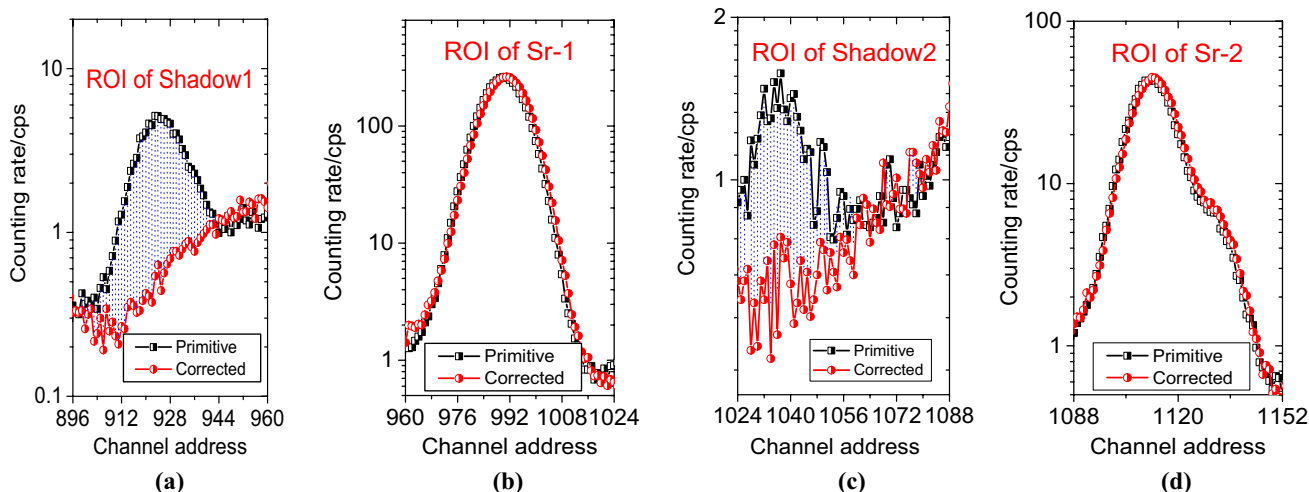


Fig. 11 (Color online) Correction effect of shadow peak of strontium element characteristic peak



**Fig. 12** (Color online) Spectral comparison of different ROI before and after correction. **a** ROI of shadow1, **b** ROI of Sr-1, **c** ROI of shadow2, **d** ROI of Sr-2

characteristic peak ROI before and after the spectral peak correction, as shown in Eq. (14).

To quantify the correction effect of the CNN-LSTM model on the X-ray spectra of the measured iron ore samples, two indicators, the correction ratio  $R_c$  and the efficiency ratio  $R_e$ , were introduced. The correction ratio represents the proportion of the increment of the peak area to the peak area of the corrected characteristic peak ROI, and the efficiency ratio represents the proportion of the increment of the peak area of the two characteristic peak ROI after calling the model to the peak area of the two shadow peak ROI. The calculation formulas are given by Eqs. (15) and (16), respectively. Ten measurements were performed on the iron ore samples, and the measurement results were analyzed, as shown in Table 3.

$$R_c = \frac{S_{\text{peakloss1}} + S_{\text{peakloss2}}}{\sum_{i=960}^{1024} \text{Count}_{i-\text{corrected}} + \sum_{i=1088}^{1152} \text{Count}_{i-\text{corrected}}} \times 100\% \tag{15}$$

$$R_e = \frac{S_{\text{peakloss1}} + S_{\text{peakloss2}}}{S_{\text{shadow2}} + S_{\text{shadow2}}} \times 100\% \tag{16}$$

From the comparison of the measurement results, it can be seen that the corrected X-ray spectrum obtained using the CNN-LSTM model to predict the pulse height has two typical features. First, the peak area of the characteristic peak ROI was improved compared with the original spectrum, and the standard deviation of the multiple measurement results was also significantly reduced. Second, the peak area of the shadow peak area was significantly reduced. According to the two characteristics above and

**Table 3** Details of measurement results

Times	Original spectrum				Corrected spectrum				$R_c(\%)$	$R_e(\%)$
	$S_{\text{shadow1}}$	$S_{\text{Sr-1}}$	$S_{\text{shadow2}}$	$S_{\text{Sr-2}}$	$S_{\text{shadow1}}$	$S_{\text{Sr-1}}$	$S_{\text{shadow2}}$	$S_{\text{Sr-2}}$		
1	126.91	4231.21	68.29	822.07	48.58	4304.88	52.51	837.99	1.74	95.20%
2	127.26	4289.77	76.46	855.27	51.32	4341.12	49.44	877.24	1.41	71.21
3	121.42	4271.21	68.42	831.42	53.39	4317.29	55.14	844.31	1.14	72.52
4	121.67	4199.54	72.51	819.33	44.25	4286.53	41.25	824.59	1.80	84.88
5	127.36	4229.79	76.79	827.93	49.87	4314.18	45.87	841.02	1.89	89.92
6	131.48	4262.17	73.87	831.85	45.66	4351.29	43.73	847.34	2.01	90.21
7	127.75	4281.21	70.41	846.83	49.32	4358.91	46.83	846.88	1.49	76.22
8	124.61	4209.39	76.82	820.44	55.38	4291.27	52.31	826.72	1.72	94.05
9	134.59	4191.29	75.11	790.12	42.26	4276.72	38.22	832.29	2.50	98.75
10	117.75	4213.71	71.33	811.58	47.11	4277.11	51.61	824.28	1.49	84.22
Avg	126.08	4237.93	73.00	825.68	48.71	4311.93	47.69	840.27	1.72	86.27
STD	4.70	33.79	3.15	17.16	3.85	28.66	5.19	14.94	-	-

the energy conservation theorem, it can be inferred that the peak area reduced in the shadow peak area should theoretically be corrected to the characteristic peak ROI, and the correction effect can be evaluated by  $R_e$  defined above. Table 3 shows that approximately 86.27% of the peak area reduced in the shadow 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 is approximately 1.72%, which is of great significance for X-ray spectroscopy and elemental content analyses.

## 4 Conclusion

In this study, we trained a CNN–LSTM model for peak correction using X-ray fluorescence spectroscopy. The model processed randomly generated distorted pulses. To improve training efficiency, the model was divided into two parts: feature extraction was performed on the data using a CNN and pulse amplitude estimation was performed using the LSTM network. In the simulation, the relative errors in the amplitude estimation of pulse sequences with different degrees of distortion were obtained using triangular shaping, CNN–LSTM, and LSTM models. As a result, for severely distorted pulses, the relative error of the CNN–LSTM model in estimating the pulse parameters was reduced by 14.35% compared to that of the triangular shaping algorithm; for slightly distorted pulses, the relative error of the CNN–LSTM model was reduced by 0.33%.

During the experiment, FAST-SDD was used to perform X-ray measurements on the iron ore samples. The measured pulse sequence was saved offline as the model input, and the pulse amplitude output of the model was analyzed for multichannel pulse height, resulting in an X-ray energy spectrum corrected for shadow peaks. Meanwhile, the original energy spectrum obtained without calling the model was used as a reference spectrum for comparison with the corrected spectrum. The results indicate that the proposed model successfully predicts the heights of the measured pulse sequences. To further validate the performance of the model for shadow peak correction, ten measurements of iron ore samples showed that the peak area of the shadow peak ROI decreased by approximately 86.27%, which can be corrected to the characteristic peak ROI, and the corrected peak area accounted for approximately 1.72% of the characteristic peak ROI. This is of great significance for X-ray spectroscopy and elemental analysis.

**Acknowledgements** The authors thank the Sichuan XSTAR Technology of M&C Co. Ltd. for providing the experimental platforms.

**Author contributions** All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Lin Tang, Xing-Ke Ma, Kai-Bo Shi, Yeng-Chai Soh, and Hong-Tao Shen. 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 final manuscript.

**Data availability** The data that support the findings of this study are openly available in Science Data Bank at <https://www.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 competing interests.

## References

1. M.N. Nader, D.E.B. Fleming, Assessment of alternative methods for analyzing X-ray fluorescence spectra. *Appl. Radiat. Isotopes*. **146**, 133–138 (2019). <https://doi.org/10.1016/j.apradiso.2019.01.033>
2. H.X. Wu, R.J. Di, Y.J. Liu et al., Application on straight-line shaping method for energy spectrum measurement in TXRF spectrometer based on SDD detector. *Spectrosc. Spectr. Anal.* **41**(7), 2148–2152 (2021). (in Chinese)
3. V.T. Jordanov, Unfolding-synthesis technique for digital pulse processing. Part 1: unfolding. *Nucl. Instrum. Meth. A* **805**, 63–71 (2015). <https://doi.org/10.1016/j.nima.2015.07.040>
4. C. Zhao, H. Xiong, Y. Liu et al., A new digital filter based on sinusoidal function for gamma spectroscopy. *Nucl. Instrum. Meth. A* **944**, 162582 (2019). <https://doi.org/10.1016/j.nima.2019.162582>
5. X.D. Zhong, L. Chen, B.C. Wang et al., A spectrometer with baseline correction and fast pulse pile-up rejection for prompt gamma neutron activation analysis technology. *Rev. Sci. Instrum.* **89**(12), 123504 (2018). <https://doi.org/10.1063/1.5049517>
6. Y.L. Song, F.Q. Zhou, Y. Li et al., Methods for obtaining characteristic  $\gamma$ -ray net peak count from interlaced overlap peak in HPG  $\gamma$ -ray spectrometer system. *Nucl. Sci. Tech.* **30**, 11 (2019). <https://doi.org/10.1007/s41365-018-0525-7>
7. M. Kafae, M.M. Goodarzi, Pile-up correction in spectroscopic signals using regularized sparse reconstruction. *IEEE Trans. Nucl. Sci.* **67**, 858–862 (2020). <https://doi.org/10.1109/TNS.2020.2985104>
8. M.Y. Kantor, A.V. Sidorov, Detection of true Gaussian shaped pulses at high count rates. *J. Instrum.* **15**, P06015 (2020). <https://doi.org/10.1088/1748-0221/15/06/P06015>
9. S. Usman, A. Patil, Radiation detector deadtime and pile up: a review of the status of science. *Nucl. Eng. Technol.* **50**(10), 1006–1016 (2018). <https://doi.org/10.1016/j.net.2018.06.014>
10. 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>
11. 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**, 1760–17678 (2018). <https://doi.org/10.1107/S160057751801411X>
12. X. Peng, H. Zhu, X. Zhou et al., ECG signals segmentation using deep spatiotemporal feature fusion U-Net for QRS complexes and R-peak detection. *IEEE Trans. Instrum. Meas.* **72**, 1–12 (2023). <https://doi.org/10.1109/TIM.2023.3241997>

13. D. Zoni, J. Flich, W. Fornaciari, CUTBUF: buffer management and router design for traffic mixing in VNET-Based NoCs. *IEEE Trans. Parall. Distr.* **27**(6), 1603–1616 (2016). <https://doi.org/10.1109/TPDS.2015.2468716>
14. L. Tang, S. Zhou, Y. Li et al., Application of multi-head attention mechanism with embedded positional encoding in amplitude estimation of stacked pulses. *Nucl. Tech.* **46**(9), 090505 (2023). [https://doi.org/10.11889/j.0253-3219.2023.hjs.46.090505\(inChinese\)](https://doi.org/10.11889/j.0253-3219.2023.hjs.46.090505(inChinese))
15. N.T. Cheng, J. Fu, D.J. Chen et al., An antibody-free liver cancer screening approach based on nanoplasmonics biosensing chips via spectrum-based deep learning. *NanoImpact* **21**, 100296 (2021). <https://doi.org/10.1016/j.impact.2021.100296>
16. Y.F. Chen, W. Han, Harmonic current control strategy of DC distribution network based on deep learning algorithm. *Energy Rep.* **8**, 13066–13075 (2022). <https://doi.org/10.1016/j.egyr.2022.09.071>
17. N.A.S. Alwan, Z.M. Zahir, Deep learning for robust adaptive inverse control of nonlinear dynamic systems: improved settling time with an autoencoder. *Sensors* **22**(16), 5935–5935 (2022). <https://doi.org/10.3390/s22165935>
18. T.H. Wang, Y. Lei, Z. Tian et al., Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy. *J. Med. Imaging* **6**(4), 043504 (2019). <https://doi.org/10.1117/1.JMI.6.4.043504>
19. H.R. Liu, Z. Zuo, P. Li et al., Anti-noise performance of the pulse coupled neural network applied in discrimination of neutron and gamma-ray. *Nucl. Sci. Tech.* **33**, 75 (2022). <https://doi.org/10.1007/s41365-022-01054-6>
20. J. Tamura, K. Futatsukawa, Y. Kondo et al., Measurement of H<sup>0</sup> particles generated by residual gas stripping in the Japan proton accelerator research complex linac. *Nucl. Instrum. Meth. A* **1049**, 168033 (2023). <https://doi.org/10.1016/j.nima.2023.168033>
21. M. Gomez-Fernandez, W.K. Wong, A. Tokuhiko et al., Isotope identification using deep learning: an explanation. *Nucl. Instrum. Meth. A* **988**, 164925 (2021). <https://doi.org/10.1016/j.nima.2020.164925>
22. C.H. Kim, S. Ahn, K.Y. Chae et al., Noise signal identification in time projection chamber data using deep learning model. *Nucl. Instrum. Meth. A* **1048**, 168025 (2023). <https://doi.org/10.1016/j.nima.2023.168025>
23. M. Touch, D.P. Clark, W. Barber et al., A neural network-based method for spectral distortion correction in photon counting X-ray CT. *Phys. Med. Biol.* **61**(16), 6132–6153 (2016). <https://doi.org/10.1088/0031-9155/61/16/6132>
24. A. Regadío, L. Esteban, S. Sánchez-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>
25. B. Jeon, S. Lim, E. Lee et al., Deep learning-based pulse height estimation for separation of pile-up pulses from NaI(Tl) detector. *IEEE Trans. Nucl. Sci.* **69**(6), 1344–1351 (2022). <https://doi.org/10.1109/TNS.2021.3140050>
26. X.Y. Guo, L. Zhang, Y.X. Xing, Study on analytical noise propagation in convolutional neural network methods used in computed tomography imaging. *Nucl. Sci. Tech.* **33**, 77 (2022). <https://doi.org/10.1007/s41365-022-01057-3>
27. X.K. Ma, H.Q. Huang, X. Ji et al., Estimation of trapezoidal-shaped overlapping nuclear pulse parameters based on a deep learning CNN–LSTM Model. *J. Synchrotron Radiat.* **28**, 910–918 (2021). <https://doi.org/10.1107/S1600577521003441>
28. 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>
29. A. Graves, A.R. Mohamed, G. Hinton. Speech recognition with deep recurrent neural networks: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645–6649(2013). <https://doi.org/10.1109/ICASSP.2013.6638947>

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>  · Xing-Ke Ma<sup>5</sup> · Kai-Bo Shi<sup>1</sup> · Yeng-Chai Soh<sup>4</sup> · Hong-Tao Shen<sup>2</sup>

✉ Xing-Ke Ma  
maxingke@stu.cdut.edu.cn

✉ 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 and Application, Anhui University, Hefei 230039, China

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

<sup>5</sup> College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Chengdu 610059, China