The Development of Advanced Deep Learning-Based EoR Signal Separation Techniques

S. Pradeep, C. V. P. R. Prasad, and Ch Ruchitha

Abstract The weak EoR signal is submerged in the strong foreground radiation interference. The classic the foreground removal methods assume that the foreground spectrum is smooth, but the complex instrumental effect will affect the spectral structure, resulting in the failure to accurately detect the signal. In this paper, a deep learning network called CNN-LSTM model is proposed to separate the EoR signal to improve computing resource utilization for massive observational data. Based on the simulation data of SKA1-LOW, a CNN-LSTM fusion model is constructed to reconstruct the EoR signal. Experimental results show that compared with the traditional methods including polynomial fitting and continuous wavelet transform, the EoR signals detected by the proposed deep learning model have better quantitative evaluation indexes of SNR and Pearson correlation coefficient. This property provides a new way to explore the research field of EoR.

Keywords OR signal · Deep learning · Signal separation algorithm

1 Introduction

The 21-cm neutral hydrogen line records the evolution stage of the universe from neutral hydrogen filled to ionized hydrogen. Studying the cosmic EoR signals of great significance for exploring the epoch of EoR (EoR) and the formation of the first generation of celestial bodies in the evolution of the universe. The radiation intensity of EoR signal is lower than 10 mK, while the brightness temperature of foreground radiation signal is about 4–5 orders of magnitude higher, and the power SNR reaches

S. Pradeep (✉) · C. V. P. R. Prasad
Department of CSE, Malla Reddy Engineering College for Women (UGC AUTONOMOUS), Maisammaguda, Hyderabad, Telangana, India
e-mail: pradeep.sunkari87@gmail.com

C. Ruchitha
Department of IT, Malla Reddy Engineering College for Women (UGC AUTONOMOUS), Maisammaguda, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

R. Malhotra et al. (eds.), *High Performance Computing, Smart Devices and Networks*, Lecture Notes in Electrical Engineering 1087,
https://doi.org/10.1007/978-981-99-6690-5_4

about -50 dB. The weak EoR signal is submerged in the strong foreground radiation interference, containing a large amount of noise caused by instrument effect. It indicates that accurate identification and separation of foreground interference is the key to solve the problem of detecting EoR signal in image space and power spectrum space.

At present, some methods can be used to detect the EoR signal. The traditional EoR signal separation methods mainly include foreground subtraction and foreground avoidance. The foreground subtraction method is to separate the EoR information from the strong foreground disturbance in image space or Fourier space by using the different characteristics of foreground contamination and EoR information displayed on the spectrum. Based on such ideas, V. Jellić et al. constructed a parametric model to describe the spectral variation of the foreground, so as to fit and deduct the pollution in the simulated radio sky map, which has a significant effect on removing the large-scale foreground interference [1]. After that, Adrian Liu used weighted polynomials to fit the information content of each pixel. On the basis of maintaining the original effect, Adrian Liu processed the small-scale foreground components more effectively [2]. Anna Bonaldi applied the correlation component analysis (CCA) method to the simulated data of square kilometer array, effectively clearing the foreground and noise in the observation data within the exploration frequency range (100–200 MHz) [3]. However, the selection of fitting parameters makes the experimental results fluctuate greatly, and both over fitting and under fitting will cause the real features of the foreground to be unable to be accurately captured, thus affecting the detection effect of the EoR signal [3]. The Wp smoothing method proposed by Harker [4] avoids the fitting residuals being polluted by the power leakage in the foreground and makes full use of spectral features to achieve separation. Gu uses continuous wavelet transform to greatly reduce the calculation amount on the basis of ensuring the same reconstruction effect as Wp smoothing method [5].

Such methods require a high degree of smoothness in the foreground radiation spectrum to clearly distinguish the reionized signal. In many cases, the foreground deduction method may mistake some large-scale EoR signals as pollution and deduct them, which makes the fitted signals differ greatly from the input signals. Ian H assumed that the foreground interference existed in a region of the two-dimensional power spectrum and processed its power spectrum to avoid the foreground contaminated area to extract EoR signals. This foreground avoidance method reduced the deviation of the foreground deduction method, but it would cause the information of the observed data outside the EoR window to be ignored in the processing [6]. The work of Patil et al. discussed the influence of various factors on the EoR power spectrum by calibrating residual sidelobe noise, nonlinear effects, and gain errors caused by noise and ionosphere at the baseline [7].

In addition, the processing of massive data collected by large interferometric arrays such as SKA1-LOW is also a key bottleneck [8]. With high sensitivity and resolution, the equipment can detect signals with complex morphology and extremely high radiation intensity range, which provides more observation information for the detection of EoR signals.

Literature [9–11] shows its outstanding performance in galaxy cluster identification and classification, pulsar search, and other aspects. These applications indicate that deep learning has the potential to extract weak signals from highly time-varying data and can be used to solve the problem of EoR signal separation.

In this paper, with the aid of OSKAR [https://github.com/OxfordSKA/OSKAR] and WSClean [https://sourceforge.net/projects/wsclean/] simulation with instrument effect data radio sky, the deep learning algorithm is used to construct the CNN-LSTM fusion model to separate the reionized signal from the complex foreground pollution, and it is compared with the traditional polynomial fitting and continuous wavelet transform method, which provides a new idea for detecting the reionized signal.

1.1 EoR Signal Separation Approach

At present, a number of methods have been proposed to try to separate EoR signals from foreground interference, including the traditional foreground deduction method, foreground avoidance method, and the newly developed crossover study of deep learning and EoR signal. In the study of EoR signals, the most classical methods are polynomial fitting [12] and continuous wavelet transform [5]. In fact, the separation algorithm has different treatment effects on the pollution components with different physical sources and morphological characteristics.

1.2 Foreground Removal Method

Di Matteo first uses the angle fluctuation of the 21-cm spectral line to separate the foreground contamination, but the effective signal is submerged by various foreground radiation and difficult to separate. Later, Zaldarriaga et al. shifted their focus to the frequency correlation of foreground and used the mutual correlation as deduction, which evolved into the line of sight (LOS) method [13], that is, in image space or UV space (Fourier space). Based on the distinct spectral features of foreground radiation and EoR signal, smooth foreground components are identified and subtracted for each pixel along the frequency dimension (i.e., line of sight direction). In the following research, Chapman et al. [14] divided it into the method that adopts functional form (parameterized) for foreground signal and the method that slightly liberates this form (non-parameterized), aiming to find the form of smooth foreground function along frequency for each line of sight and subtract it from the total signal. Leaving behind noise, fitting errors, and residuals of the cosmological signal.

2 Polynomial Fitting

Early parametric methods used polynomials to fit directly to the measured data in the frequency or logarithmic frequency domain. The usual way of polynomial fitting is to fit the total observed spectrum along the line of sight with a smooth function, such as a polynomial of order N:

$$\log T_{b,\mathrm{fg}}(v) = a_0 + \sum_{i=1}^{n} a^i \log v^i \tag{1}$$

The order n is set as 2, and Jelic adjusts the order n to 3 and discusses the influence of polynomial order on the fitting effect. If the order of polynomial is too small, the foreground fitting will be insufficient, and the fitting residuals will affect the EoR signal. If the order of the polynomial is too large, the EoR signal is over fitted.

After stripping the radio source with very large brightness temperature in the frequency domain space, the foreground interference is deducted along the line of sight according to the characteristic that the foreground component has smoother spectrum than the EoR signal. In other words, the logarithm of the total signal intensity I_{total} of the observed frequency received by pixel V_i in the line of sight direction gives the function.

$$Y_i = \lg(I_{\mathrm{Total}}(V_i)) \tag{2}$$

The intensity of foreground component was simulated by multi-order polynomial fitting to the logarithmic signal I_{fore},

$$I_{\mathrm{fore}}(V_i) \approx a_0 + a_1 \lg(v_i) + a_2 \lg^2(v_i) \tag{3}$$

EoR signal radiation and instrument noise can be expressed as the difference between the overall radiation intensity and the fitted polynomial signal intensity.

According to the different angular power spectrum of the signal, the reconstructed signal is obtained by subtracting the instrument noise $I_{\mathrm{EoR}\prime}$

$$I_g(I_{\mathrm{Total}}(V_i)) \approx a_0 + a_1 \lg(v_i) + a_2 \lg^2(v_i) + I_{\mathrm{EOR}}\left(v^i\right) \tag{4}$$

$$I_{\mathrm{EOR}\prime}(v_i) \approx I_{\mathrm{Total}}(v_i) - 10^{a_0+a1} \lg(v_i) + a_2 \lg^2(v_i) \tag{5}$$

Fitting the foreground signal with various curve functions will have different influences on the extraction of EoR signal, such as exponential function, Fourier function, Gaussian function, and power exponential function. Chapman et al. showed in their research work that the EoR signal extracted by linear fitting the foreground signal is closer to the simulated signal [14]. In the work of Bonaldi [15]. and Brown, the idea of

polynomial fitting was introduced into Fourier domain correlation component analysis, and second-order statistics were used to estimate the spectrum of foreground components from data [13]. For this study, we will adopt the conclusion of Jelic et al. and use third-order polynomials to fit in log space.

2.1 Continuous Wavelet Transform

Non-parametric methods do not directly assume that foreground spectra should conform to a particular parametric model, or any spatial structure about them, but make full use of the different spectral characteristics of foreground radiation and EoR signals to achieve the separation of the two. For example, typical methods are Wp smoothing, independent component analysis, ICA, generalized morphological component analysis (GMCA), and continuous wavelet transform (CWT).

In this paper, we use the continuous wavelet transform method outlined by Harker et al., and according to the inverse CWT definition proposed by Daubechies [16],

$$h(t) = \frac{2}{C_\psi} \int_0^\infty \left[\int_{-\infty}^{\infty} W_{x,\psi}(\tau, s) \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s} \right) \mathrm{d}\tau \right] \frac{\mathrm{d}s}{s^2} \tag{6}$$

where $h(t)$ represents the real space signal to be transformed, $\psi(t)$ called the mother wavelet function, and τ, s represents the real space and scale exponent of the wavelet coefficient $W_{x,\psi}$. Prospect in the study of EoR, radiation, and EoR signal distribution of significant coefficients is different, and the significant coefficients of smooth prospects mainly depend on the data boundary discontinuity,

$$\psi_s(t) = \frac{\psi\left(\frac{t}{s}\right)}{\sqrt{|s|}} \tag{7}$$

where $T_d(v)$ represents the wavelet coefficient of the total radiation signal, δ_c is the Dirac delta function, and $v_{\min}$ represents the lower frequency limit. The wavelet coefficients of the total signal are filtered through a mask. Since the wavelet transform is actually the cross correlation between $\psi_s(t)$ and the real space signal $h(t)$, according to the cross correlation theorem, the real space signal can be effectively calculated in the Fourier space, and this equation is used to reconstruct the filtered EoR signal.

$$h(t) = \frac{2}{C_\psi} \int_0^\infty F^{-1}\left\{ F\left\{ W_{x,\psi} \right\} \cdot \psi_s \right\}(t, s) \frac{\mathrm{d}s}{s^2}. \tag{8}$$

Harker introduced the Wp smoothing method into the EoR detection experiment and used the physical expected values of foreground components for separation. This method uses a smoothing function to fit a set of observations of a frequency channel. Machler first studied the boundary value problem to give a smooth function formula;

Gu et al. further studied the problem and gave a new algorithm. Although a general smoothness of the foreground is assumed to make the method a good application, the method includes a smoothness parameter that allows the user to control the harshness of this smoothness condition to allow for deviations from the prior smoothness.

In addition, Chapman et al. applied FastICA to EoR simulations as an independent component analysis technique, which assumes that foreground components are statistically independent in order to model them. Emma Chapman et al., using the component separation method, generalized morphological component analysis (GMCA) [14] has successfully conducted separation experiments on simulated LOFAR data.

2.2 Foreground Avoidance Method

The advantage of foreground deduction method is that it can retain all the information of EoR signal, but the disadvantage is that it may not be able to deduct foreground pollution accurately or some large-scale EoR signals are mistaken for foreground and deducted, resulting in a certain degree of deviation in the results. In order to effectively avoid the possible deviation caused by foreground pollution on EoR detection results, foreground avoidance method has been paid more attention and studied in recent years.

In the study of Chapman et al. [17], it is shown that by deducting the foreground by modeling or by analyzing the specific part of the EoR delay power spectrum that is not affected by the foreground, there is an "EoR window" in the two-dimensional power spectrum theoretically because most of the instrumental color effect is confined to the wedge. The effectiveness of this method mainly depends on the instrument response and the smoothness of the observed sky [18].

Liu et al. proposed a mathematical formalism for describing wedges [19] that allows maximizing the range of the EoR window. Several methods have also been developed to estimate the covariance of prospects, which can then be incorporated into the power spectrum estimator. However, these foreground avoidance or suppression methods have the disadvantage of substantially reducing the sensitivity of the instrument, as they reduce the number of detectable patterns. Moreover, Rajesh Mondal et al. [20] Future observation studies using the upcoming SKA-LOW measure the prospect of EoR signals and in this work quantitatively addresses the impact of prospects on map delectability predictions by avoiding signals contained within the planar foreground wedge.

2.3 Deep Learning Method

In recent years, some literatures have explored the idea of foreground removal based on machine learning, and several attempts have been made to improve the performance and accuracy of neural networks for the task of detecting EoR signals, with

varying degrees of success. Samuel Gagnon-Hartman et al. described and tested the implementation of a U-Net architecture [21] with the aim of using information from Fourier patterns to identify ionized regions in wedge-filtered images and to be able to reconstruct their shape, size, and location in the image. In the work of Li et al. [22].

In addition, Shimabukuro et al. [23] introduced artificial neural networks (ANNs) to constrain astrophysical parameters from the 21-cm power spectrum to rigorously estimate the 21-cm power spectrum from the interferogram while mitigating foreground contamination, thus improving sensitivity. In La Plante's work, the convolution neural network (CNN) is used to further constrain the cosmological parameters during the EoR period. By using several supervised learning methods to improve prediction accuracy, including neural networks, kernel regression, or ridge regression, this work compares the performance of these methods using a large training set of SKA simulated 21-cm power spectra.

In contrast, in our work, in order to study the EoR signal separation problem, we treat foreground radiation as strong noise and EoR signal as effective signal, so the problem is transformed into the identification problem of EoR signal: By removing the noise (i.e., foreground radiation) in the total radiation signal (i.e., the sum of EoR signal and foreground radiation), the signal (i.e., EoR signal) is obtained, and the separation of EoR signal and foreground radiation is realized.

3 Methods

3.1 CNN BiLSTM Network Architecture

We built a neural network to eliminate the foreground radiation seen in the instrument beam pattern and separate the weak EoR signal. The effect will be demonstrated with the help of a radio sky simulating SKA observations. According to the astrophysical theory, we preprocessed the OSKAR simulated visibility data to exclude the influence of instrument effect as much as possible and then standardized the data.

3.2 Classic Convolutional Neural Network

The essence of the problem of detecting the EoR signal is to separate the foreground contamination and extract the EoR signal from the total radiation signal. One-dimensional convolutional neural network (CNN) can extract a variety of abstract features from data and has a good performance effect in the separation of mono sound sources. The core of neural network is to extract features layer by layer through convolution operation,

$$X_j^{(l)} = f\left(\sum_{i \in M_j} X_j^{(l-1)} * K_{ij}^{(b)} + b_j^{(b)}\right) \tag{9}$$

where l the number of is convolutional layers; $X_j^{(l)}$ is the first feature map of the layer, $K_{ij}^{(b)}$ is the convolution kernel, and b is the bias term. $*$ stands for convolution operator.

3.2.1 Long Short-Term Memory

The bidirectional long short-term memory network (LSTM) derived from recurrent neural network (RNN) has a strong ability to extract weak features from highly time-varying data. This model mainly adds three gating structures of hidden layers to the original recurrent neural network. The calculation method is as follows:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right), \tag{10}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{11}$$

$$\widetilde{C}_t = \tan h\left(W_C \cdot \left[h_{t-1}, x_t\right] + b_C\right) \tag{12}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t, \tag{13}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{14}$$

$$h_t = o_t * \tan h(C_t) \tag{15}$$

The parameters in the above formula are updated through the internal unit structure of LSTM, where x_t represents the input value at the current unit time, $h_{t\text{-}1}$ represents the output value at the previous unit time, $\widetilde{C}$ represents the state information of candidate unit, and C_t represents the internal state information of current unit. The input gate i determines the part of the input information that can be reserved for C_t; the forgetting gate f decides to save the part of the unit state C_{t-1} at the previous moment according to the data feature information of the attention feature; the output gate o determines the part of the hidden state at the moment $t-1$ that can be transferred to the current state. Its unit structure is shown in Fig. 1.

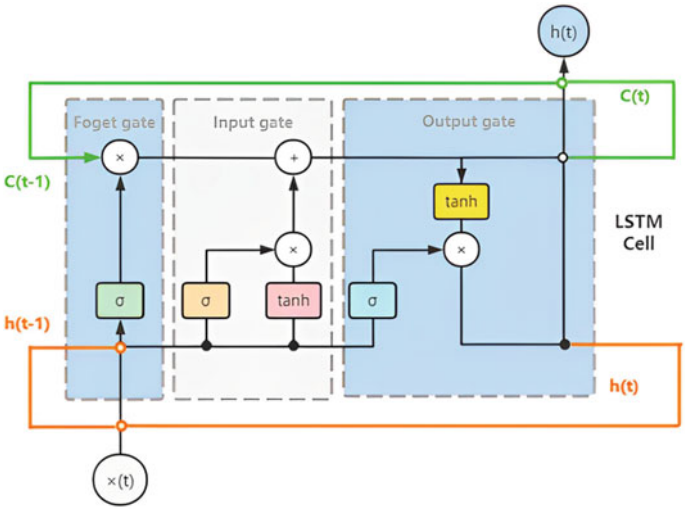

Fig. 1 LSTM internal structure diagram

3.2.2 Fusion Module

The analysis can predict that the two networks have the ability to process astronomical data and separate EoR signals. Since the signal-to-noise ratio of the signal to be measured is very low, we take advantage of the characteristics of CNN and LSTM and combine the advantages of the two frameworks to construct a new CNN-BiLSTM network.

In the test phase, we constructed two kinds of fusion deep neural networks, which were vertically and horizontally combined with CNN and LSTM, respectively, and finally selected the vertical cascade structure with better performance as the model basis. In this model, LSTM receives the output of CNN and extracts features in a deeper way. The architecture of the CNN-BiLSTM model is shown in Fig. 2, followed by a detailed overview of the framework.

The local characteristics of the input data used in this paper, the model could reflect the key information of phase spectrum change point, convolution neural network (CNN) with multiple local feature extraction of convolution layer to the data, because each convolution filter layer contains a fixed size, and quantity greatly reduced, in the use of the limited weight of back propagation algorithm to update the parameters, After data preprocessing, each pixel in the observation area is listed in the dimension of observation frequency as the input of three one-dimensional convolutional layers, in which the first convolutional layer contains 32 filters of size 3×1, and the second and third convolutional layers contain 16 filters of size 3×1. The eigenvector is obtained by operation according to the equation v_i,

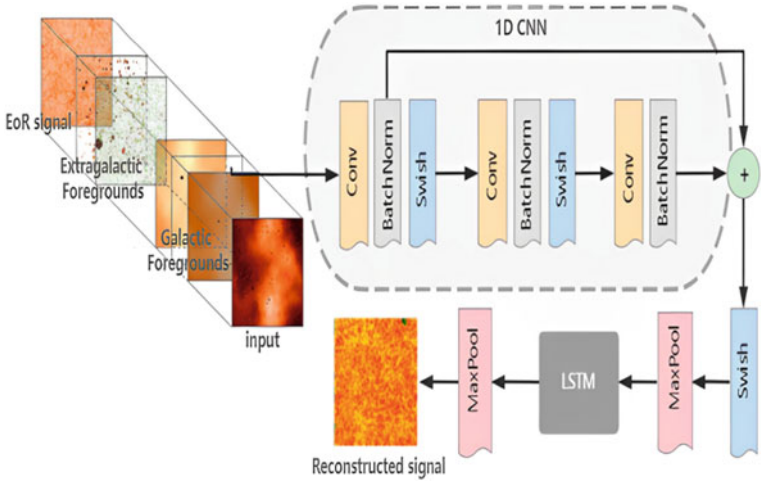

Fig. 2 Network architecture diagram

$$v_i^{(l)} = \phi^{(l)}\left(\sum_{j=1}^{m_{l-1}} v_j^{(l-1)} * W_i^{(l)} + b_i^{(l)}\right), \quad i = 1, 2, \ldots, m_1 \tag{16}$$

where l represents the sequence number of convolutional layer, m_1 represents the total number of filters in layer l, φ represents the activation function of this layer, $W_i^{(l)}$ and $b_i^{(l)}$, respectively, represent the weight parameter and bias term of layer l, and set $\left\{v_i^{(l)}\right\}$; $i = 1, 2, \ldots, m_1$ represents the sum of feature vectors obtained from all filters in layer l.

3.2.3 Activation Function

By comparing and testing the effects of different activation functions, Swish function is finally determined as the activation function of network convolutional layer.

$$f(x) = x * \sigma(\beta x) \tag{17}$$

where σ stands for sigmoid function. Sigmoid function has saturation property, and introducing β can train parameters to avoid gradient disappearance. When β takes different values, the function image is as follows (Fig. 3).

Swish function is essentially a smooth function between linear function and ReLU function. Since the sparse processing forced by ReLU will reduce the effective capacity of the model, when the gradient is 0 when $x < 0$, it often leads to the problem of neuron death. The Swish function can correct this problem and adjust

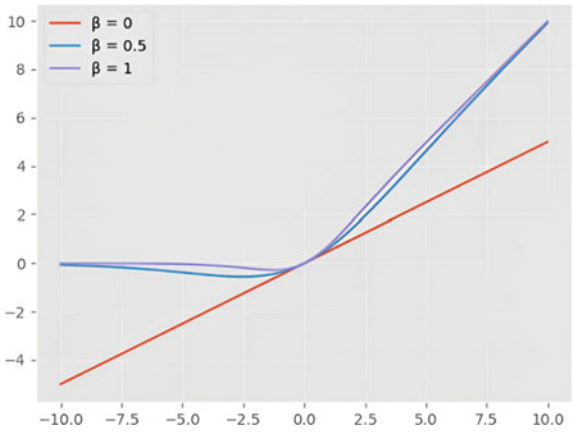

Fig. 3 Swish function curve

the training parameters suitable for the model so that the network can achieve higher performance than the standard ReLU.

After maximum pooling, the feature vectors are passed through the LSTM layer containing hidden units, and the state activation function of TANH and the gate activation function of sigmoid are used. In addition, we add an appropriate dropout layer to the CNN model to discard 20% of random features and weaken the joint adaptability between neuronal nodes, thus avoiding overtraining of LSTM sequence data. At the same time, Adam algorithm is used to optimize the objective function to further improve the performance of the model.

3.3 Model Training

The CNN-LSTM network architecture is trained and tested. Firstly, uniform initialization was used to set all parameters of the network to random values. The initial learning rate was set as 0.003, and the batch size was set as 125. The total radiation spectrum data of pixels were input into the network model one by one for convolution operation. After 100 rounds of iterative training, EoR signals were separated from the network model. The loss function used to guide parameter tuning has been tested and decided to quantify the difference between the EoR signal obtained from the model and the EoR signal input to the network using cross-entropy, which is defined as

$$C = -\frac{1}{n}\sum_{x}[y \ln a + (1-y)\ln(1-a)] \tag{18}$$

In formula (19), x represents the sample, y represents the actual label, a represents the predicted output, and n represents the total number of samples. This loss function

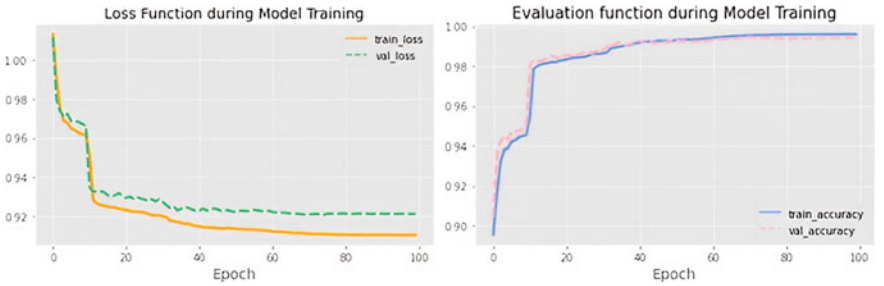

Fig. 4 Model training process: curve of loss function (left panel) and Evaluation function (right panel)

avoids the problem that the weight of the square loss function is updated too slowly. When the error is large, the weight update speed is improved, and when the error is small, the weight update speed is slowed down, so that the parameter update strategy can be adjusted in time. At the same time, the evaluation function is set, and the accuracy is calculated to evaluate the performance of the current training model.

Figure 4 shows the changes of the loss function and evaluation function with the training process, taking the training period as the variable. The final network accuracy is 0.972, indicating that the network has good performance.

4 Experiments

4.1 Datasets

The observed data used in the experiment are from the OSKAR simulation data based on the SKA1-LOW site layout and the EoR simulation data published by the Evolution of 21 cm Structure project. Taking (R.A.,Dec.) = (3°,−27°) as the pointing center, the first image cube containing foreground and EoR signals is simulated in the frequency band range of 154–162 MHz. For the EoR signal, the data should be standardized after eliminating the outliers. We divided 70% of the data points in the first set of image cubes into training sets, and the remaining pixels were used as validation. All the data in the second image cube are used as the test set to test the network performance and reconstruct the complete image of EoR signal. Figure 5 shows the entire framework of the preprocessing process.

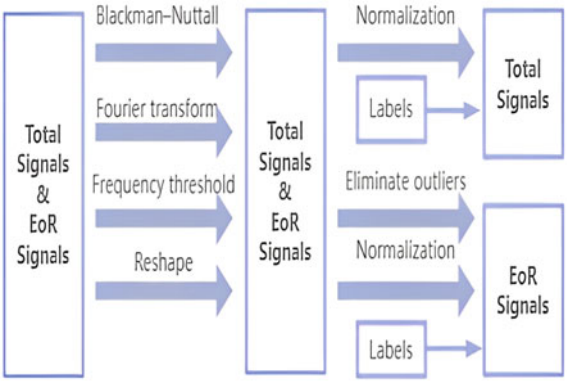

Fig. 5 Data preprocessing pipeline

4.2 Model Performance

In the study of EoR, the CLEAN algorithm has limitations on the processing effect of faint diffuse radiation, so we took advantage of the "Faintgalaxies" Lightcone image cube to generate sky maps by WSClean. As shown in Fig. 6. The EoR image shows the density distribution corresponding to the red shift, while in the superposed low-frequency sky, the EoR signal with relatively low bright temperature is completely covered by the intense foreground radiation.

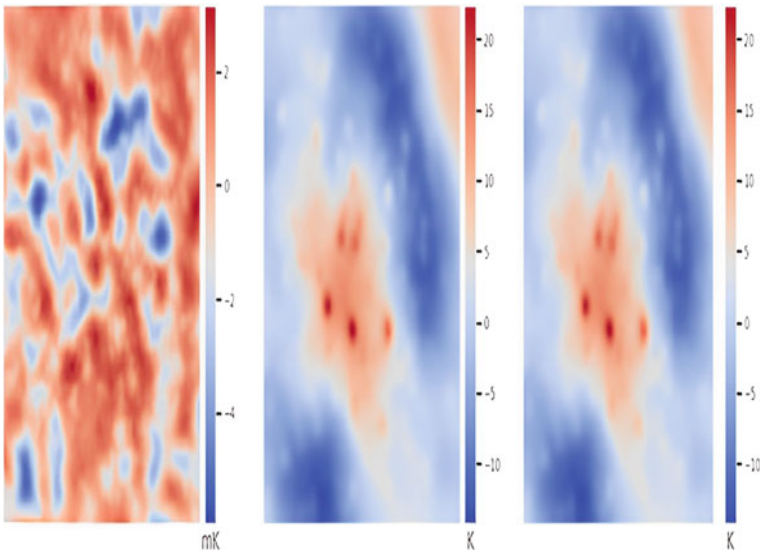

Fig. 6 Simulated images of the EoR signal (left panel), the foreground emission (middle panel), and sky map (right panel) at 160 MHz

The EoR signal after model processing was imaged, and the difference of EoR signal before and after reconstruction was compared. In Fig. 7, for detecting signals and reconstruction, overall space structure and the outline is almost the same, choose the same color display standardized after radiation intensity distribution. This is because there are fine small-scale ripples structures (small ripples) in the foreground Fig. 8, there are narrow bands in the power spectrum, making the fitting effect deviation, optimize the model to get a more accurate fitting effect is our next improvement direction.

Figure 9 shows the variation curve of the magnetic field intensity of the received foreground radiation signal, EoR signal, and total signal simulated by a randomly selected sky pixel as a function of frequency. It can be seen that the EoR signal presents a sawtooth spectral structure, but there is a gap of 4–5 orders of magnitude

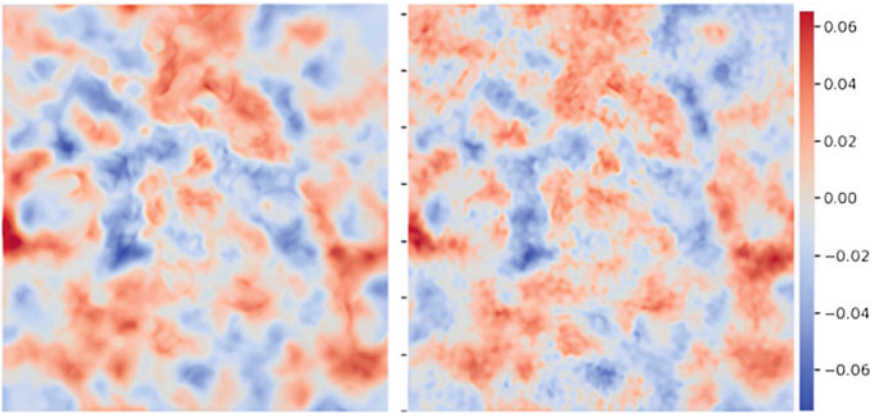

Fig. 7 Comparison between reconstructed EoR signal (left panel) and reconstructed EoR signal (right panel)

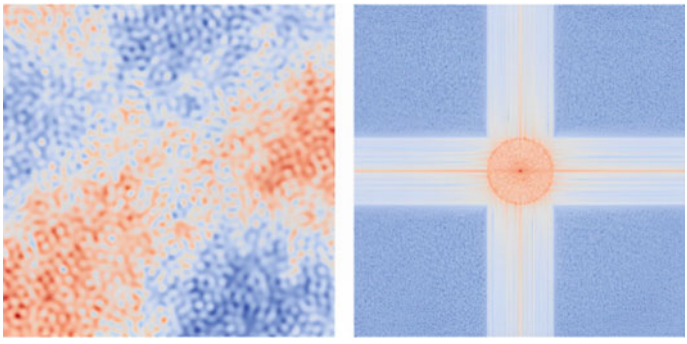

Fig. 8 Small ripples of original foreground signal (left panel) and bright stripe in the 2D power spectrum (right panel)

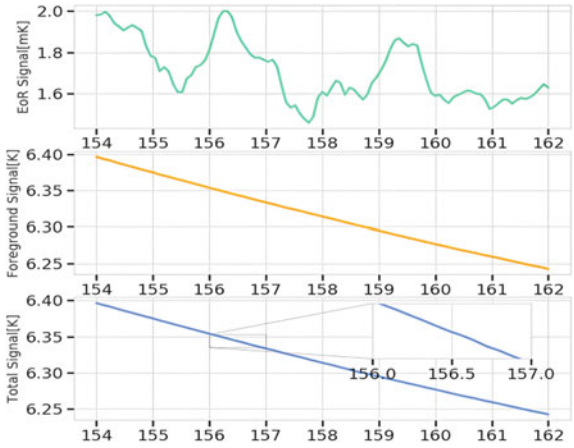

Fig. 9 EoR radiation signal foreground radiation signaling on total radiation signal

between its brightness temperature and that of foreground component. The EoR signal cannot be visually distinguished in the total signal.

4.3 Result Comparison

In order to verify the difference between the proposed method and the traditional method, qualitative and quantitative comparisons of the effects of the two evaluation methods were used to compare the separation of EoR signals.

A. *Spectral Analysis*

In the observed frequency band, the spectrum characteristics of the reionized signal obtained by the continuous wavelet transform method, polynomial fitting, and deep learning model are compared with the original signal. In order to emphasize the change effect, the radiation signal is normalized. Can be seen from the Figs. 10, 11, and 12 model is presented by this paper get the signal is compared with the other two traditional methods to get more close to the input signals of ionization signal, again shows that under the complex beam effect can still treat signal fluctuation of smaller amplitude and scale of reconstruction, get more complete ionization signal spectrum structure again.

B. *Quantitative Analysis*

In order to quantify the similarity between the reconstructed reionized signal and the input signal, Pearson correlation coefficient and the signal-to-noise ratio (R_{sn}) of the separated signal and the original signal are introduced as quantitative analysis indicators to measure the reconstruction effect, so as to better reflect the numerical differences of the results obtained by different methods. Combined with the understanding of the signal magnetic field strength, the signal-to-noise ratio index (R_{sn})

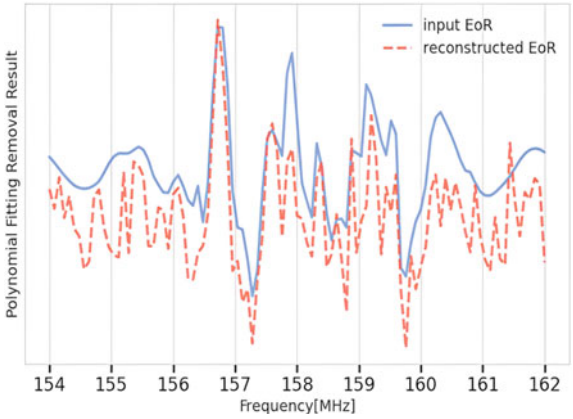

Fig. 10 Result of polynomial fitting method

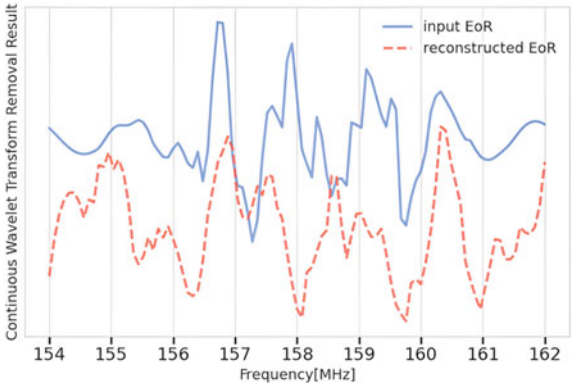

Fig. 11 Result of continuous wavelet transform method

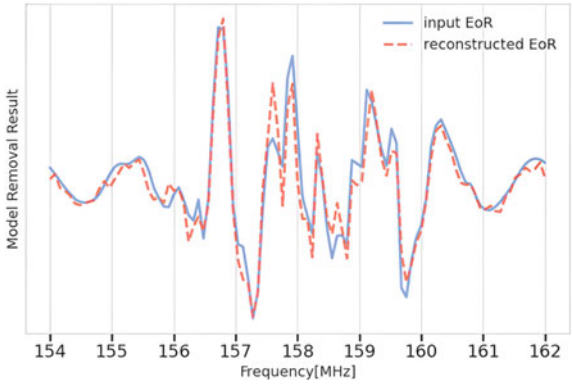

Fig. 12 Result of CNN-LSTM model

Table 1 Comparison of the performance of EoR signal reconstruction between traditional methods and deep learning method

Index	Rsn/dB	Pearson
Polynomial fitting method	9.352	0.477
Continuous wavelet transform method	13.340	0.316
CNN-LSTM model	17.231	0.813

and Pearson correlation coefficient were defined as follows

$$R_{\mathrm{sn}} = 10 \lg \frac{E(|r_{\mathrm{eor}}|)^2}{E(|x_{\mathrm{eor}} - r_{\mathrm{eor}}|)^2}, \tag{19}$$

$$(\mathbf{r}_{\mathrm{eor}}, \mathbf{x}_{\mathrm{eor}}) = \frac{\sum_{j=1}^{n} \left(r_{\mathrm{eor},j} - \overline{r}_{\mathrm{eor}}\right)\left(x_{\mathrm{eor},j} - \overline{x}_{\mathrm{eor}}\right)}{\sqrt{\sum_{j=1}^{n} \left(r_{\mathrm{eor},j} - \overline{r}_{\mathrm{eor}}\right)^2 \sum_{j=1}^{n} \left(x_{\mathrm{eor},j} - \overline{x}_{\mathrm{eor}}\right)^2}}. \tag{20}$$

where r_{eor} is reconstruction EoR signal, x_{eor} is original input EoR signal, and n is the length of the signal. Pearson's correlation coefficient is close to 1, indicating that the more accurate the reconstructed EoR signal is, the better the performance of the processing method is. Table 1 shows the statistical results of the SNR separation performance index, reflecting the different performance effects of polynomial fitting method with order 4 and continuous wavelet transform method and deep learning model for the same test data.

These results fully demonstrate that the trained deep learning model can effectively separate the spectrum of foreground radiation and EoR signals, so as to accurately detect EoR signals.

5 Conclusions

The existing EoR signal separation algorithms are not good for signal reconstruction when dealing with the foreground components with complex morphological changes in massive observation data. This paper first introduces the traditional polynomial fitting and wavelet transform method and then discusses the deep learning model from the theoretical aspect. After that, it is verified by experiments that the CNN-LSTM model is better than the traditional EoR signal separation algorithm in removing foreground components, and the SNR and Pearson correlation coefficient are used as the evaluation indexes of quantitative analysis to evaluate the separation results of different methods. The research can provide a new reference idea for EoR signal separation, which can save computing resources such as memory and time, and reduce the requirement of hardware. For further improving the reconstruction effect, there are still many key problems to be solved, such as the complex instrument effect

introduced in the observation and the size of the imaging sky area, which will affect the detection. In the next work, we will study the influence of wide-field imaging of SKA1-LOW on small-scale radiation celestial signals and explore the dispersion degree and magnetic field intensity changes of point signals at different phases far from the zenith. Since the experiment in this paper is conducted on simulated data, the processing effect of observation data from other telescopes needs to be further explored.

References

1. Jelić V, Zaroubi S, Labropoulos P et al (2010) Foreground simulations for the LOFAR—epoch of EoR Experiment. Mon Not R Astron Soc 389(3):1319–1335
2. Liu A, Tegmark ME, Bowman J et al (2009) An improved method for 21cm foreground removal. Mon Not R Astron Soc 398(1):401–406
3. Anna B, Brown ML (2015) Foreground removal for square kilometre array observations of the epoch of EoR with the correlated component analysis. Mon Not R Astron Soc 2:1973–1983
4. Harker G, Zaroubi S, Bernardi G et al (2009) Non-parametric foreground subtraction for 21-cm epoch of EoR experiments. Mon Not R Astron Soc 397(2):1138–1152
5. Gu J, Xu H, Wang J et al (2013) The application of continuous wavelet transform based foreground subtraction method in 21 cm sky surveys. Astrophys J 773(1):773–779
6. Ian H, Emma C, Pritchard JR et al (2020) Comparing foreground removal techniques for recovery of the LOFAR-EoR 21cm power spectrum. Mon Not R Astron Soc 2:2264–2277
7. Patil AH, Yatawatta S, Koopmans L et al (2017) Upper limits on the 21 cm epoch of EoR power spectrum from one night with LOFAR. Astrophys J 838(1):65–81
8. Eloy DLA, Trott CM, Wayth RB et al (2018) Spectral performance of SKALA antennas I: mitigating spectral artefacts in SKA1-LOW 21-cm cosmology experiments. Mon Not R Astron Soc 3:2662–2671
9. Lukic V, Brüggen M (2017) Galaxy classifications with deep learning. Proc Inter Astronom Union 12(S325):217–220
10. Burger B, Mattia V, Matthew P et al (2021) CNN architecture comparison for radio galaxy classification. Mon Not R Astron Soc 503(2):1828–1846
11. Lin H, Li X, Zeng Q (2020) Pulsar candidate sifting using multi-input convolution neural networks. Astrophys J 899(2):104–116
12. Wang X, Tegmark M, Santos MG et al (2006) 21 cm tomography with foregrounds. Astrophys J 650(2):529–539
13. Loeb A, Zaldarriaga M (2005) The small-scale power spectrum of cold dark matter. Physical Rev D 71(10):467–470
14. Chapman E, Abdalla FB, Harker G et al (2012) Foreground removal using fastica: a showcase of LOFAR-EoR. Mon Not R Astron Soc 423(3):2518–2532
15. Bonaldi A, Battye RA, Brown ML (2014) Cosmological constraints from Sunyaev-Zeldovich cluster counts: an approach to account for missing redshifts. Astrophys J 786(2):86–93
16. Ingrid D et al (1992) Sets of matrices all infinite products of which converge. Linear Alge Appl 161(15):227–263
17. Chapman BE, Harker A et al (2015) Cosmic dawn and epoch of EoR foreground removal with the SKA. Physics 9525(12):9009–9014
18. Parsons AR, Backer DC, Bradley RF et al (2010) The precision array for probing the epoch of re-ionization: eight station results. Astron J 139:1468–1480
19. Liu A, Parsons AR, Trott CM (2014) Epoch of EoR window. I. Mathematical formalism. Physical Rev, D. Particles, Fields, Gravitation, Cosmo 90(2):18–23

20. Mondal R, Bharadwaj S, Majumdar S (2016) Statistics of the epoch of EoR 21-cm signal–I. Power spectrum error-covariance. Monthly Notices Royal Astronomical Soc 456(2):1936–1947
21. Gagnon-Hartman S, Cui Y, Liu A et al (2021) Recovering the lost wedge modes in 21-cm foregrounds. arXiv e-prints, arXiv: 2102.08382
22. Li W, Xu H, Ma Z et al (2019) Separating the EoR signal with a convolutional denoising autoencoder: a deep-learning-based method. Mon Not R Astron Soc 485(2):2628–2637
23. Shimabukuro H, Ichiki K, Inoue S et al (2014) Probing small-scale cosmological fluctuations with the 21 cm forest: effects of neutrino mass, running spectral index and warm dark matter. Physical Rev D 90(8):83–86