

A review of EEG noise reduction technology based on DCCRN-LSTM

Weitao Lu

Fuzhou University, Maynooth International Engineering College, Fuzhou, 350108, China

WEITAO.LU.2022@MUMAIL.IE

Abstract. Electroencephalography (EEG) represents a pivotal technology in Brain-Computer Interface (BCI) research, offering insights into brain activity through non-invasive methodologies. However, EEG signals are frequently contaminated by noise sources, necessitating the application of denoising techniques to improve signal quality. This paper reviews a denoising model combining Long Short-Term Memory (LSTM) networks and Deep Complex Convolutional Recurrent Network (DCCRN) for effective EEG noise reduction. The combination of LSTM and DCCRN offers a powerful approach to mitigate diverse noise components in EEG data, as both networks are adept at addressing gradient vanishing and explosion issues, respectively, and handling complex-valued operations. The DCCRN-LSTM model has the potential to enhance the quality of EEG data, thereby improving the accuracy and robustness of subsequent signal processing tasks. The application of DCCRN, which is predominantly used in speech denoising, introduces new methodologies to the realm of EEG signal processing, and it has the potential to be a significant contributor to the future field of EEG noise reduction.

Keywords: EEG, LSTM, DCCRN, Complex network, denoise.

1. Introduction

The brain computer interface (BCI) concept was first proposed in 1973. This technology involves the establishment of a direct connection between a computer and the brain of a human or animal, creating a communication and control system that bypasses traditional muscle and nerve pathways [1]. Following its initial proposal, BCI technology underwent extensive development without any breakthroughs for an extended period. Nevertheless, it was not until the 1990s that significant progress was made, driven by advances in computer science, neurology, and brain research. This technology has made significant advances in several fields, including clinical medicine, neurorehabilitation, art, and entertainment. BCI systems can be classified according to the methods employed for acquiring signals, namely non-invasive and invasive systems. Non-invasive systems offer a resolution of approximately 0.05 seconds and 1 millimeter, while invasive systems provide a much finer resolution of about 0.003 seconds and 0.05 millimeters [2], as detailed in Sections 2.1 and 2.2. However, in the development of BCI technology for many years, the processing of noise and artefacts has been a major difficulty, this paper will use literature analysis and other methods to study Noise and artefacts in EEG signals are processed using the DCCRN-LSTM algorithm to preserve privacy and enhance valid information in EEGs, which can reduce the

interference of noise for the effective information in EEG and provide a potential noise reduction method for the field of EEG analysis. This study focuses on EEG noise reduction technology based on DCCRN-LSTM. For evaluating related articles the criterion was the relevance of EEG and DCCRN-LSTM to this paper, which are the key areas of concern in this thesis.

2. Classification of BCI systems based on different ways of acquiring signals

2.1. Invasive Method

Invasive BCI involves the placement of electrodes directly into the brain. This approach offers higher temporal and spatial resolution than non-invasive methods. However, it typically necessitates a craniotomy, which involves the removal of a portion of the skull, the placement of the electrodes or probes onto the cerebral cortex, and then the restoration of the skull. Consequently, invasive BCI is more expensive and carries greater risks, including the potential for infection and the degradation of recorded signals over time.

2.2. Non-Invasive Method

Typical non-invasive technologies include Near-infrared Spectroscopy (NIRS), Magnetoencephalography (MEG), Functional Magnetic Resonance Imaging (fMRI), and EEG. While non-invasive brain-computer interfaces (BCIs) offer lower temporal and spatial resolution than the invasive method, they do not require surgical implantation. Instead, they can be easily implemented using a headband or hat, making them cost-effective and relatively safe. By placing electrodes on the scalp and applying conductive gel, or using headgear to secure the electrodes and directly record neural signals, these systems provide the benefits of affordability, ease of operation, and safety [3]. Consequently, non-invasive BCI technology is widely adopted and utilized. The Electroencephalogram (EEG), is a technique that captures electrical impulses produced by brain neurons by placing electrodes and other electronic components on the scalp [4]. However, EEG signals are frequently susceptible to a variety of disturbances, including those resulting from human physiological processes or non-physiological sources, such as mechanical equipment. The nonlinear, random, non-correlated, and non-Gaussian characteristics of EEG signals render the effectiveness of traditional wavelet transform or filter methods in eliminating these interferences severely limited. In order to address this challenge, deep learning-based methods have demonstrated significant potential in the realm of noise reduction. This article presents a novel approach that employs a combination of DCCRN and LSTM to mitigate EEG noise.

3. Key algorithms for noise reduction and EEG enhancement

3.1. LSTM

Long Short-Term Memory (LSTM) is a framework derived from RNN (Recurrent Neural Networks) [5]. In traditional RNNs, connections are typically established between hidden layers, with the input to a hidden layer including both the output from the input layer and the output from the hidden layer at the previous time step. While RNNs offer advantages in analysing time series data, they frequently encounter issues such as long-term dependency problems, gradient vanishing, and gradient explosion [6]. LSTM effectively addresses these challenges by preserving long-term sequence information and mitigating the issues of gradient vanishing and explosion. Compared to the AutoRegressive Integrated Moving Average (ARIMA) time series model, LSTM is more adept at handling complex time series prediction tasks, particularly those involving long-term dependencies.

The fundamental gate mechanism of LSTM comprises the input, forget, and output gates. The introduction of the forget gate enables LSTM to retain information over extended periods. Upon receipt of a novel input, LSTM combines the input with the output from the preceding time step to generate a distinct vector. Subsequently, the aforementioned vector is subjected to a sigmoid neural layer. If the value is close to 0, the component information has been “forgotten.” Conversely, a value close to 1 indicates the retention of complete memory. This principle is exemplified by the following formula.

$$f_t = \sigma(W_f \cdot [y_{t-1}, x_t] + a_f) \quad (1)$$

Where f_t represents the decision vector of the forget gate, σ , W_f , respectively represent the sigmoid activation function and the weight of the forget gate of dimension $d_c \times (d_x + d_h)$; y_{t-1} , x_t are the output at time t-1 and the input at time t; a_f represents the offset of the forget gate.

The output value of the current LSTM is:

$$y_t = \tanh(C_t * o_t) \quad (2)$$

C_t represents the new cell state, o_t represents the decision vector of the output gate.

$$i_t = \sigma(W_i[y_{t-1}, x_t] + a_i) \quad (3)$$

$$\bar{C}_t = \tanh(W_c[y_{t-1}, x_t] + a_c) \quad (4)$$

In equations (3) and (4), i_t represents the decision vector of the input gate, and $\bar{C}_t$ represents the candidate information.

$$C_t = \bar{C}_{t-1} * f_t + \bar{C}_t * i_t \quad (5)$$

$$o_t = \sigma(W_o[y_{t-1}, x_t] + a_o) \quad (6)$$

In the context of EEG denoising, the LSTM model can learn features from long time series data. During training, the LSTM model utilises the learned features and patterns to effectively denoise the EEG signals. The LSTM model can predict and remove noise components from the EEG data, thereby improving the signal-to-noise ratio.

3.2. DCCRN

Machine learning models are typically classified into time series models and regression models. In the context previously outlined, the LSTM model is employed due to its capacity to learn features from long time series data. Consequently, this section employs a time series model for modelling purposes. The Deep Complex Convolution Recurrent Network (DCCRN) is a network model that combines the structures of Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) to address the issue of time series dependencies. It is distinctive in that it incorporates complex convolutional layers and LSTM layers. The complex network is capable of capturing the correlation between amplitude and phase angle through complex multiplication, thereby enabling a more comprehensive representation of signal characteristics [7].

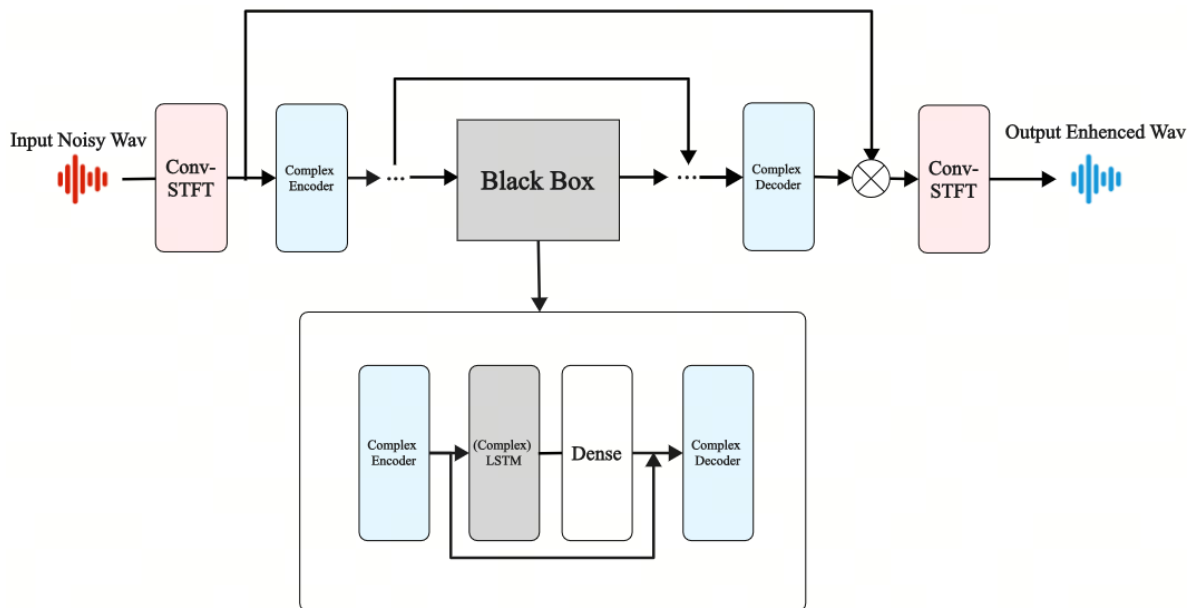


Figure 1. DCCRN Flowchart

The encoder/decoder block comprises a complex two-dimensional convolutional neural network (Conv2d), batch normalization, and a real-valued parametric rectified linear unit (PreLU) [8]. The complex-valued convolution filter K can be represented as $K = K_r + jK_i$, where K_r is the real part of the complex convolution kernel, and K_i is the imaginary part of the complex convolution kernel. The input complex matrix M can be expressed as $M = M_r + jM_i$. In conclusion, the complex output expression is derived through the operation $X \otimes W$, as illustrated in Equation (7).

$$F_{out} = (M_r * K_r - M_i * K_i) + j(M_r * K_i + M_i * K_r) \quad (7)$$

In formula (7), F_{out} denotes the output feature of one complex layer.

According to the real part M_r and imaginary part M_i of the complex input, the output F_{out} of the complex LSTM can be expressed as:

$$F_{rr} = \text{LSTM}_r(M_r); F_{ir} = \text{LSTM}_r(M_i) \quad (8)$$

$$F_{ri} = \text{LSTM}_i(M_r); F_{ii} = \text{LSTM}_i(M_i) \quad (9)$$

$$F_{out} = (F_{rr} - F_{ii}) + j(F_{ri} + F_{ir}) \quad (10)$$

LSTM_r and LSTM_i represent the real and imaginary parts in the traditional LSTM model, while F_{ri} represents the result calculated by LSTM_i and input M_r .

The original design of DCCRN was intended for speech enhancement tasks, with a particular focus on phase-aware speech enhancement. Although DCCRN was not initially designed for EEG noise reduction, its deep learning capabilities and complex number processing offer a promising potential solution for this application. Nevertheless, further research and experimental validation are required to ascertain the efficacy of this approach in EEG noise reduction. In particular, the application of complex number processing to EEG data, along with the adjustment and optimisation of the network structure, necessitates further extensive exploration and investigation.

3.3. Section summary

As a recurrent neural network structure, LSTM is particularly adept at processing time series data and capturing long-term dependencies. In EEG signal processing, LSTM is an effective tool for modelling complex patterns in time series. This is crucial for noise reduction tasks due to the presence of helpful information and noise in EEG signals over extended time scales. In contrast, DCCRN has been developed for the specific purpose of processing complex-valued data. It offers the advantages of low training parameters and computational costs and does not require prior knowledge. In the context of EEG signal processing, the use of complex-valued operations enables the capture of both amplitude and phase information, which is of particular importance in retaining key features within the signal and in the removal of noise. The integration of LSTM and DCCRN within an algorithm has the potential to enhance noise reduction efficiency, resulting in reduced computing time and resource consumption. The long-term dependency capture capability of LSTM, when combined with the complex value operation processing capability of DCCRN, results in a synergistic effect, enabling the algorithm to more effectively identify and separate noise components in EEG signals, thereby improving the overall noise reduction effect. The principles, advantages, and disadvantages of various algorithms are presented in Table 1.

Table 1. Principles, advantages and disadvantages of various algorithms

Algorithm Name	Principle	Advantage	Disadvantage
LSTM	The control of the forget gate, input gate, and output gate enables the processing of long-term dependencies in time series data.	Capable of resisting the phenomenon of vanishing and exploding gradients.	Not applicable to spatial data

Table 1. (continued).

DCCRN	Optimizing Scale-Invariant Source-to-Noise Ratio (SI-SNR) loss by combining complex CNN, complex batch normalization layer and complex LSTM	Excellent performance and high parameter efficiency	High computational complexity
Regression methods	By establishing a relationship model between features and original signals, the error between predicted values and actual values is minimized, and the coefficient parameters in the linear regression model are estimated.	Simplify the model and reduce the amount of calculation	Need to get a better noise source signal
Wavelet Transform	Decompose the signal into sub-signals of different frequencies and filter out noise by selecting appropriate wavelet basis functions and threshold processing methods [9]	The time-frequency characteristics are optimal.	High computational complexity
Principal component analysis	Eigenvalue implementation based on covariance matrix [10]	Computationally efficient, no additional information required	Difficulty distinguishing interference when drift potential is similar to EEG signal
Adaptive filtering	The weights are iteratively adjusted according to the optimization algorithm to quantify the noise in the main input	High adaptability	Requires additional reference input

4. Conclusion

This paper reviews EEG denoising technology based on the DCCRN-LSTM model. The LSTM's ability to mitigate gradient explosion and vanish with the DCCRN's strengths in complex-valued operations enables the effective processing of various noise components in EEG signals. The successful application of this technology can significantly enhance the quality of EEG data, thereby improving the accuracy and robustness of signal processing. Although DCCRN is currently primarily employed in speech denoising, it introduces novel concepts and methodologies to EEG signal processing. However, the method mentioned in this paper still has some shortcomings. The DCCRN-LSTM algorithm is not widely used in the field of EEG, so the method may have poor performance for certain types of signals and noise in EEG, which needs to be optimised and improved after further use and observation of the model, and there is also some room for improvement in the large computational complexity of the model. Future research will focus on optimising the model structure and parameters of the DCCRN-LSTM model to improve its performance in EEG noise reduction, followed by further investigation of the possibility of combining it with other algorithms to enhance the noise reduction effect. In addition, the application of this algorithm in other fields needs to be explored to increase its wide range of applications. By addressing these limitations of the DCCRN-LSTM model, and for future focused directions, the model has great potential in the field of EEG noise reduction, which will improve the processing capability of EEG signals and provide a potential solution to the noise reduction problem for more directions.

References

- [1] Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*, 2(1), 157-180.
- [2] Fernando Nicolas-Alonso, L., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, 12(2), 1211-1279.
- [3] Veena, N., & Anitha, N. (2020). A review of non-invasive BCI devices. *Int. J. Biomed. Eng. Technol*, 34(3), 205-233.
- [4] Adams, E. J., Scott, M. E., Amarante, M., Ramírez, C. A., Rowley, S. J., Noble, K. G., & Troller-Renfree, S. V. (2024). Fostering inclusion in EEG measures of pediatric brain activity. *npj Science of Learning*, 9(1), 27.
- [5] Pamungkas, Y., Wibawa, A. D., & Rais, Y. (2022, December). Classification of Emotions (Positive-Negative) Based on EEG Statistical Features using RNN, LSTM, and Bi-LSTM Algorithms. In *2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)* (pp. 275-280). IEEE.
- [6] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306.
- [7] Yang, Q., & Liu, S. (2021, December). DCCRN-SUBNET: A DCCRN and SUBNET Fusion Model for Speech Enhancement. In *2021 7th International Conference on Computer and Communications (ICCC)* (pp. 525-529). IEEE.
- [8] Hu, Y., Liu, Y., Lv, S., Xing, M., Zhang, S., Fu, Y., ... & Xie, L. (2020). DCCRN: Deep complex convolution recurrent network for phase-aware speech enhancement. *arXiv preprint arXiv:2008.00264*.
- [9] Subasi, A., Tuncer, T., Dogan, S., Tanko, D., & Sakoglu, U. (2021). EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier. *Biomedical Signal Processing and Control*, 68, 102648.
- [10] Buzzell, G. A., Niu, Y., Aviyente, S., & Bernat, E. (2022). A practical introduction to EEG time-frequency principal components analysis (TF-PCA). *Developmental cognitive neuroscience*, 55, 101114.