EEG AUTHENTICATION BASED ON DEEP LEARNING OF TRIPLET LOSS

J. Cui[*], *L. Su*[*], *R. Wei*[*†], *G. Li*[‡], *H. Hu*[*], *X. Dang*[§]

Abstract: As a novel biometric characteristic, the electroencephalogram (EEG) is used for biometric authentication. To solve the challenge of efficiently growing the number of classifications in traditional classification networks and to increase the practicality of engineering, this paper proposes an authentication approach for EEG data based on an attention mechanism and a triplet loss function. The method begins by feeding EEG signals into a deep convolutional network, maps them to 512-dimensional Euclidean space using a long short-term memory network combined with an attention mechanism, and obtains feature vectors for EEG signals with identity information; it then adjusts the network parameters using a triplet loss function, such that the Euclidean distance between feature vectors of similar signals decreases while the distance between signals of a different type increases. Finally, the recognition method is evaluated using publicly available EEG data sets. The experimental results suggest that the method maintains the recognition rate while effectively expanding the classifications of the model, hence thus boosting the practicability of EEG authentication.

Key words: *electroencephalogram, brain-computer interface, deep learning, attentional mechanisms, triplet loss*

Received: April 25, 2022 **DOI:** 10.14311/NNW.2022.32.016
Revised and accepted: October 30, 2022

1. Introduction

With the rapid advancement of information technology, authentication of personal identifiable information is critical to network information security. Biometric authentication technologies that are often used include those based on voiceprint [1], face [2], and fingerprint [3]. However, these recognition systems have a number of disadvantages, including the ease with which they can be forged, their inability to function following damage, and their lack of in vivo detection. To address these

*Jun Cui; Lei Su; Hongwei Hu; School of Life Sciences, Tiangong University, Tianjin, China
†Ran Wei – Corresponding author; School of Precision Instrument and Opto-electronics Engineering,Tianjin University, Tianjin, China; School of Life Sciences, Tiangong University, Tianjin, China, E-mail: ranwei_tgu@163.com
‡Guangxu Li; School of Electronics and Information Engineering, Tiangong University, Tianjin, China
§Xin Dang; College of Intelligence and Computing, Tianjin University, Tianjin, China; School of Computer Science and Technology, Tiangong University, Tianjin, China

shortcomings, EEG signals have been proposed as a new biometric authentication medium that satisfies the fundamental requirements of biometric authentication [4].

The electroencephalogram is produced by inducing an electric field on the subject's scalp, the characteristics of which are determined by the firing of spatially arranged cortical pyramidal neurons, and this brain activity is typically classified into five distinct oscillatory rhythms [5], namely delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma ($\geq$ 30 Hz). The characteristics of EEG signals are unique and enduring aspects of the brain, and compared to conventional biometric identification, EEG signals offer more visible benefits [6]: (i) High concealment: EEG signals are generated within the brain, necessitating the need for a specialized acquisition instrument; (ii) Liveness detection: creation of EEG signals must ensure owner activity in the area where they are placed, and EEG signals will vanish when the owner dies; (iii) Difficult to harm [7]: Compared to fingerprints, faces, and other biometric features that are exposed to the outside world for an extended period of time, the EEG is contained within the head and is the most vital and well-protected organ, which ensures that the EEG signal can be used as a biometric authentication without fear of causing damage to the EEG or allowing it to be identified. As a result, in various application scenarios that require a high level of security, such as military, financial industry, and other critical situations, EEG offers numerous advantages that enable it to be used in scenarios that require a high level of security.

With the rapid advancement of brain-computer interface (BCI) technology and the application of deep learning technology in recent years, EEG signal-based authentication systems have garnered increased attention and study. Research can generally be classified into four groups based on the cognitive tasks that trigger EEG. The first is research on the identification of EEG in the resting state. Lan Ma et al. [8] used convolutional neural networks (CNN) to automatically extract and classify the best and most unique neural characteristics of individuals, using EEG data from the open resting state of the eyes (REO) and the closed resting state of the eyes (REC), and demonstrated that they achieved high recognition accuracy in a 10-level classification. Additionally, by focusing on a very low frequency band, significant interpersonal variations can be seen. Banee Bandana Das et al. [9] suggested a spatiotemporally dense architecture for EEG-based person recognition, first using a CNN to analyze raw EEG data and then using a long short-term memory network (LSTM) to handle temporal data and perform closed-eye person recognition. The recognition rates of the individuals were 99.95 % and 98 % in the closed-eye and open-eye states, respectively. The second is EEG recognition based on visual evoked potentials (VEP). Heba El-Fiqi et al. [10] advocated the use of CNN in conjunction with the raw steady-state visual evoked potential (SSVEP) for the recognition and validation of individuals. On the SSVEP dataset, the suggested technique achieved an average recognition accuracy of 96.8 %, which is 45.5 % greater than the average recognition accuracy of other methods. Hongze Zhao et al. [11] recognized people using a template matching-based identification algorithm originally designed for the detection of VEP in BCI. Experiments demonstrated that code-modulated visual evoked potentials (c-VEPs) achieved the maximum correct recognition rates when 3.15 s of VEP data were used in the sessional condition and 99.48 % when 10.5 s of VEP data were used in the cross-sessional condition. The third category is EEG recognition based on motor imagery (MI). Yingnan Sun et al. [12] suggested a novel method based on a one-dimensional convolutional long short-term memory neural network. The method was evaluated using EEG data from 109 subjects and demonstrated an extremely high average accuracy when only 16 channels of EEG signals were used. Vitaly Schetinin et al. [13] presented the group method of data handling (GMDH), which properly recognized all subjects in studies using EEG-MI benchmark data and improved recognition accuracy statistically significantly. Finally, identification of EEG using event-related potentials (ERPs). Jinan Guan et al. [14] conducted a simulated reading ERP experiment to extract and classify features from the EEG signals of numerous subjects using an unsupervised feature learning method. The experimental results indicated that: utilizing artificial neural networks, the feature vectors of five subjects were classified with an accuracy rate greater than 90 %. Additionally, Emanuemeiorana [15] proposed a deep learning method using Siamese convolutional neural network, which was carried out on a multi-session database composed of 45 subjects' EEG data. The results can be used for EEG-based biometric verification under cross-task conditions. Alyasseri Z.A.A. et al. [16] addressed the problem of EEG channel selection by using a binary version of the grey wolf optimizer and a support vector machine classifier with a radial basis function to obtain an accuracy of 94.13 % using only 23 EEG channels with 5 autoregressive coefficients of the sensor. Mohamed Benomar [17] explored the implementation of EEG-based biometrics using ResNet, Inception and EEGNet on a multi-task BED dataset for 21 subjects achieving an accuracy of 63.21 %, 70.18 % and 86.74 % for ResNet, Inception and EEGNet, respectively.

Although academics have conducted much research on EEG biometrics with encouraging results, several fundamental difficulties remain: (i) Conventional convolutional recurrent models in which long short-term memory networks are unable to highlight significant information in a long sequence, thus impairing the network's ability to learn temporal information; (ii) Existing research on traditional algorithms based on deep learning models for classification tasks. The disadvantages of commonly used loss functions, such as cross-entropy and softmax loss functions, are that they require a fixed number of categories to be defined in advance [18], and that if new categories need to be added or reduced during the training process, the entire network must be redefined and retrained, which cannot be well adapted to the training task of nonfixed categories, and the prescribed training categories have a certain risk of failure; (3) EEG signals is typically performed on 64 or 128 conductive pole channels, which is extremely uncomfortable and impractical for the EEG acquisition process, significantly reducing the user's comfort; and there is prior experimental evidence that fewer electrodes can be used for the authentication of EEG [19], suggesting that EEG authentication in realistic engineering applications can be reduced.

Therefore, in the traditional cascaded network based on a convolutional neural network and a long short-term memory network, this paper introduces the attention mechanism and the triplet loss function to extract features and classify and identify EEG signals. Experiments based on the EEG database prove that this method has good scalability while ensuring accuracy and ensures strong practicability in the engineering of EEG authentication.

2. Materials and method

In this paper, we propose a deep neural network model for EEG authentication based on the attention mechanism with long short-term memory networks and convolutional neural networks (1DCNN-ALSTM), and the overall model is depicted in Fig. 1. The model begins with the original EEG data, abstracts the original EEG data characteristics using a convolutional neural network, and then extracts the brain pattern features of the EEG signal using a long short-term memory network. Scale regularization of EEG features will be used to ensure that the final acquired brain pattern feature is a vector of features restricted to Euclidean space.

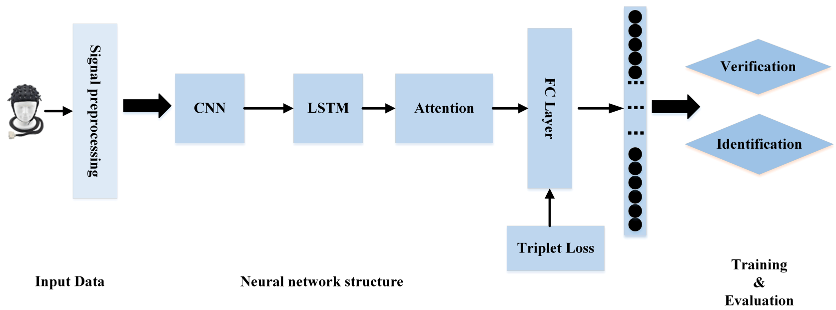

Fig. 1 *EEG authentication model.*

2.1 Description of data and pre-processing

The source of the experimental data is the publicly available Physionet EEG motor imagery data set [20]. Subjects performed different motor/imagery tasks while 64-channel EEG was recorded using the BCI2000 system [21]. The sampling frequency of the database is 160 Hz, and there are EEG data from a total of 109 subjects. Each subject performed 14 experimental runs: two one minute baseline runs (one with eyes open, one with eyes closed) and three two minute runs of each of the following four tasks: (i) A target appears on either the left or right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes. (ii) A target appears on either the left or right side of the screen. The subject imagines opening and closing the corresponding fist until the target disappears. Then the subject relaxes. (iii) A target appears at the top or bottom of the screen. The subject opens and closes either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes. (iv) A target appears at the top or bottom of the screen. The subject imagines opening and closing either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes

In order for the data to be better trained, a few minutes of EEG data from each subject are segmented so that it can have a larger number of files for the same task from the same subject. In this paper, each segment of the EEG signal is segmented into 1s EEG data. For the 1s EEG segment data, each sample can be represented as a matrix of EEG channels and sampling rates, with a total number of samples of 169246. For practical use, i.e., whether a high resolution can be achieved even without performing a specific evoked task. Therefore, the full task EEG data of each subject was used for this experiment regardless of whether this subject was performing the evoked task or not.

2.2 Neural network structure

2.2.1 One-dimensional convolution neural network

Convolution neural network has a wide range of applications in computer vision [22], and its convolutional layer uses a two-dimensional convolutional kernel to convolve image data. One-dimensional convolutional neural networks are very effective in deriving features from fixed-length segments of the entire dataset and are also known as automatic feature extractors, which automatically learn basic ordered features from the underlying raw data. Convolutional and pooling layers are commonly used for feature learning [23] and are used to extract patterns and features from the input. As for the EEG signal, which is essentially a one-dimensional sequence, a one-dimensional convolutional layer is required, and the process of one-dimensional convolutional operation is shown in Fig. 2.

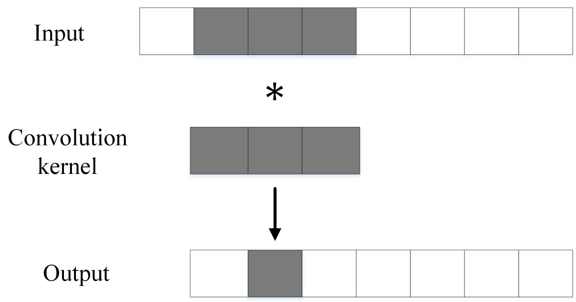

Fig. 2 *One-dimensional convolutional operation process.*

2.2.2 The long short-term memory network

The long short-term memory network is a special kind of recurrent neural network (RNN), which was proposed by Hochreiter & Schmidhuber [24] and has been applied and promoted by many people in subsequent works. It performs very well in various sequential problems, and its special structural design makes it good for solving long-dependence problems. Long short-term memory neural network can overcome the problems of "gradient disappearance" and "gradient explosion" in the backpropagation process of recurrent neural networks [25], and can effectively extract temporal features from EEG temporal signals.

The LSTM structure is mainly composed of three different gate structures: the forgetting gate, the memory gate, and the output gate. The main purpose of these three different gate structures is to control the retention and transmission of information in the LSTM, which is finally reflected in the cell state C_t and the output signal H_t. The forgetting gate consists of a sigmoid activation function and a bit multiplication operation, which is always used for information forgetting as there is always excess information in the computation of information. The forgetting gate selects the cell state C_{t-1} in which the information will be forgotten, and the activation of the forgetting gate is represented by, which is calculated using in Eq. (1).

$$F_t = \sigma(W \cdot [X_t, H_{t-1}, C_{t-1}] + b_f), \tag{1}$$

where W is the weight matrix, b_f is the offset, H_{t-1} is the output of the previous loop layer, and X_t is the input of that layer.

The memory gate consists of a sigmoid activation function and a tanh activation function, which serves the opposite function of the forgetting gate and is mainly used to remember information. It determines which of the new input information X_t and H_{t-1} retained and which is discarded. The sigmoid layer serves the same function as the forgetting gate, mainly receiving X_t and H_{t-1} as input and then outputting to a value between 0 and 1 to determine which information needs to be updated, as shown in Eq. (2). The role of the tanh layer is to integrate the input. The input X_t is integrated with H_{t-1}, and the tanh layer is used to create a new candidate vector of cell states, C_{t-1} with a value between 0 and 1, as shown in Eq. (3).

$$I_t = \sigma(W \cdot [X_t, H_{t-1}, C_{t-1}] + b_i), \tag{2}$$

$$C_{t-1} = \tanh(W \cdot [X_t, H_{t-1}, C_{t-1}] + b_c), \tag{3}$$

where W is the weight matrix, b_i and b_c are the offset, H_{t-1} is the output of the previous loop layer, and X_t is the input of that layer.

Here, the output of the forgetting gate F_t is multiplied by the cell state C_{t-1} at the previous moment to select some information for forgetting and retention, as shown in Eq. (4). The output of the memory gate is summed by the information selected by the forgetting gate to get the new cell state. C_t will continue to be passed on to the LSTM network at moment $t+1$ as the new cell state.

$$C_t = F_t \cdot C_{t-1} + I_t \cdot C_{t-1}. \tag{4}$$

After updating the cell state C_t, enter the output quantity H_t corresponding to moment t. The output gate is the state passed at moment $t-1$ and after the previous forgetting gate and memory gate, after selecting the state, integrated with the output signal at moment $t-1$ and the input signal at moment t as the current moment output signal, as shown in Eqs. (5) and (6).

$$O_t = \sigma(W \cdot [X_t, H_{t-1}, C_{t-1}] + b_o), \tag{5}$$

$$H_t = \tanh(C_t) \cdot O_t, \tag{6}$$

where b_o is the offset, O_t indicates control C_t output.

2.2.3 Attention mechanism

The attention mechanism is inspired by the human visual system and can give neural networks the ability to focus on a specific subset (input or features) [26]. Essentially, attention mechanisms filter out a small amount of valid information for focus during information processing and ignore a large amount of other invalid information to solve the information overload problem. The inability of the LSTM network to highlight important information in a long sequence weakens the network's ability to learn temporal information. The attention mechanism can compensate for this shortcoming. The attention mechanism can help the network capture stage information by introducing new trainable variables that sequentially strengthen the connection between data segments and highlight attention-focused data segments. Not all feature information is of equal importance in studies targeting EEG signals, so adding an attention mechanism to the LSTM network and assigning different weights to the output of the LTSM layer can better focus on the important feature information of the EEG signals, thus effectively improving the classification of the model [27], Eqs. (7), (8), and (9) are as follows:

$$u(t) = \tanh(W_w H_t + b_w), \tag{7}$$

$$\alpha(t) = \frac{\exp\left(u^{\mathrm{T}}(t)u_w\right)}{\sum_t \exp\left(u^{\mathrm{T}}(t)u_w\right)}, \tag{8}$$

$$S = \sum_t \alpha(t) H_t, \tag{9}$$

where W_w, u_w and b_w represent trainable weights and bias terms, respectively; $u^{\mathrm{T}}(t)$ represents the LSTM layer output of the learned features H_t; $\alpha(t)$ represents the probability distribution obtained by normalizing the weight coefficients by the softmax function, and the more important feature information in the feature information H_t extracted by the LSTM is selected and obtained by the probability distribution $\alpha(t)$, and finally the weighted sum of the feature values is obtained by the final feature information.

2.2.4 Triplet loss function

Triplet loss was first presented by Google in the FaceNet paper [28] as a loss function for solving the facial identification problem, with the goal of differentiating non-identical and very similar samples, such as two brothers. The triplet loss has three fundamental elements: three distinct embedding vectors that combine to produce a triplet. The anchor vector A, the positive vector P, and the negative vector N are the three embedding vectors. The anchor vector A serves as the reference vector, i.e., it serves as the reference vector for the following two vectors. The positive vector P represents the embedding vector of samples belonging to the same category as the anchor vector; the negative vector N represents the embedding vector of any other samples belonging to a different category. The basic concept of ternary loss is to bring the positive vector as close to the anchor vector as possible while keeping the negative vector as far away from the anchor vector as possible. This is illustrated in Fig. 3.

The margin parameter called the threshold parameter, is used to measure how close or far the positive and negative vectors are to and from the anchor vector. The triplet loss function for this parameter is shown in Eq. (10).

$$L(A, P, N) = \max(\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \alpha, 0), \tag{10}$$

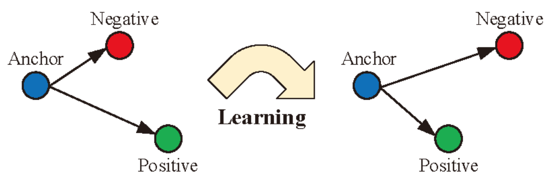

Fig. 3 *Triplet loss adjustment chart.*

where $\|f(A) - f(N)\|^2$ denotes the Euclidean distance between the anchor vector and the negative vector; $\|f(A) - f(P)\|^2$ denotes the Euclidean distance between the anchor vector and the positive vector. α is a preset very small threshold to constrain the gap between $A - N$ and $A - P$ Euclidean distance.

In the process of tuple selection, difficult samples are chosen [29] so that the triadic loss can learn more subtle features and thus better facilitate network convergence. The difficult samples include the triplet sample mining strategy, which selects the positive samples with the farthest distance from the anchor sample feature mapping and the negative samples with the closest distance from the anchor sample feature mapping under each training batch to form a difficult triad and train the neural network model. For a given anchor vector, the difference between the hard positive sample and the anchor vector is the smallest, while the difference between the hard negative sample and the anchor vector is the largest. The definitions of hard positive samples and hard negative samples are expressed in Eqs. (11) and (12), respectively, as follows.

$$hardpositive = \operatorname{argmax} \|f(A) - f(P)\|_2^2 , \tag{11}$$

$$hardnegative = \operatorname{argmin} \|f(A) - f(N)\|_2^2 , \tag{12}$$

where $\|f(A) - f(P)\|_2^2$ denotes the Euclidean distance between the anchor vector and the positive vector in the ith tuple; $\|f(A) - f(N)\|_2^2$ denotes the Euclidean distance between the anchor vector and the negative vector in the ith tuple.

2.3 Overview of proposed architecture

Fig. 4 shows 1DCNN-ALSTM network structure. The network is made up of fourteen layers, including five convolution layers, two long short-term memory network layers, four batch normalization layers, an attention mechanism layer, a full connection layer and a resize layer. The main parts of the network are the 1DCNN layer, the LSTM layer, and the attention mechanism layer

Original data is first processed by reversed convolution structure that possess 5 convolution layers, with 3×1 scale of all convolution nucleus and num of convolution nucleus for each layer is 1024, 512, 256, 128 and 64 in order. The step of each single convolution opreation for data was set to 1. When the data went through the fifth convolution layer, the shape of each EEG data was changed to 160×64. Then the whole data is reshaped from batchsize $\times 160 \times 64$ to $160 \times$ batchsize $\times 64$, and we know that 160 is about infomation of time dimension, so the input dimension of

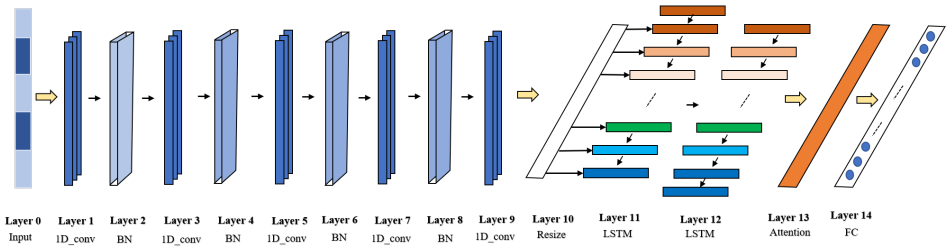

Fig. 4 *Network structure based on 1DCNN-ALSTM.*

each double-layer LSTM cell is set to 160. After going through the LSTM layer, the output goes through the attention mechanism layer. Finally, the data goes through the full connection layer and become 512D EEG feature output, which has 512 numbers to describe EEG of different individuals.

The EEG signals are mapped to a feature space constituted of 512 real numbers in this model by multiple feature extraction, and the feature vectors of the EEG identity signals are obtained. Following back propagation, the network is graded down in the direction of minimizing the triplet loss, and the network parameters are changed, as illustrated in Eqs. (13) and (14). Following repeated training, the Euclidean distance between 512-dimensional eigenvectors generated by the same identity will eventually diminish, while the distance between distinct signals will gradually increase.

$$W'_{ij} = W_{ij} - lr\frac{\partial}{\partial W_{ij}}L(W,b), \tag{13}$$

$$b'_i = b_i - lr\frac{\partial}{\partial b_i}L(W,b), \tag{14}$$

where lr is the network learning rate; W_{ij}, b network weight and bias; $L(W,b)$ is the triplet loss function.

3. Experiments and analysis of results

3.1 Experimental environment configuration

The experiment implementation environment is equipped with the Ubuntu18.04 operating system, an $Intel(R)Xeon(R)CPUE5-2680v4@2.40$ GHz CPU processor, and a $GeForceGTX1080Ti$ graphics card, and the model uses the Pytorch1.9.0 deep learning framework built on Python3.8.3. During the training phase of the experiment, each experiment was randomly separated into three groups: 80 % of the EEG signals were used to train the model, 10 % of the data were used to verify the model, and 10 % of the data were maintained for model testing. Each iterative training session in the experiment begins with the training data being combined into 64 batches. When the network reaches 200 epochs or the training loss is at standstill, the network training will be terminated.

To evaluate the classification performance of the created model, this article determines the correct recognition rate (CRR), loss values (loss), model loading time (L-time) and batch testing time (T-time) for EEG identities.

3.2 Authentication performance of the model

To highlight the benefits of the proposed 1DCNN-ALSTM model structure, we can more effectively depict the spatial information contained in the EEG in a 512-dimensional Euclidean space, resulting in a higher recognition rate. When using the same 64-channel EEG dataset and the same triplet loss function, the loss convergence of the training to the network is stopped.

The classification and identification performance indicators for the three models are shown in Tab. I, the EEG authentication model based on attention LSTM and CNN achieves a high classification rate of 99.97 %, which is 0.35 % better than the single 1DCNN model and 3.1 % higher than the single LSTM model. The loss rate of the 1DCNN-ALSTM model is the lowest in Tab. I, and while the model loading time in the test set is longer than that of the single 1DCNN and LSTM model, it is still less than 0.5 s, which meets the requirements of engineering applications. Additionally, the 1DCNN-ALSTM model achieves a minimum test time of 0.23 s for each batch, which is approximately 0.1 s faster than that of a single network. In summary, it shows that the 1DCNN-ALSTM model can outperform the standard single network model and successfully increase the accuracy of the model.

	Indicator			
Models	CRR [%]	Loss	L-time [s]	T-time [s]
1DCNN	99.61	0.0442	0.0271	0.3189
LSTM	97.87	0.1245	0.0408	0.3609
1DCNN-ALSTM	99.97	0.0230	0.3553	0.2369

Tab. I *Authentication results under 64 electrodes for different models.*

To illustrate the model's utility and to make EEG authentication realistic. Individual EEG electrodes from the BCI2000 system used in the EEG database were screened in this work for the EEG areas caused by the motor imagery EEG paradigm (region C, region P, and region CP), there were a total of 23 electrodes, as shown in Fig. 5. As a result, in this paper, the EEG data from nine electrodes in the P region, seven electrodes in the CP region, and seven electrodes in the C region were retained sequentially, and the EEG data from these electrodes were substituted into the above model for training, while the data from the remaining electrodes were erased.

The electric recognition effects of the three models based on different numbers are depicted in Tab. II. By comparison, it is discovered that as the number of electrodes is reduced, the accuracy of the model gradually decreases, but the recognition rate increases. EEG recognition rates based on the P area (9 electrodes) and the P-C-CP region (23 electrodes) are approximately equal and exceed 99 %. Although the variation in electrode count is modest compared to the CP and

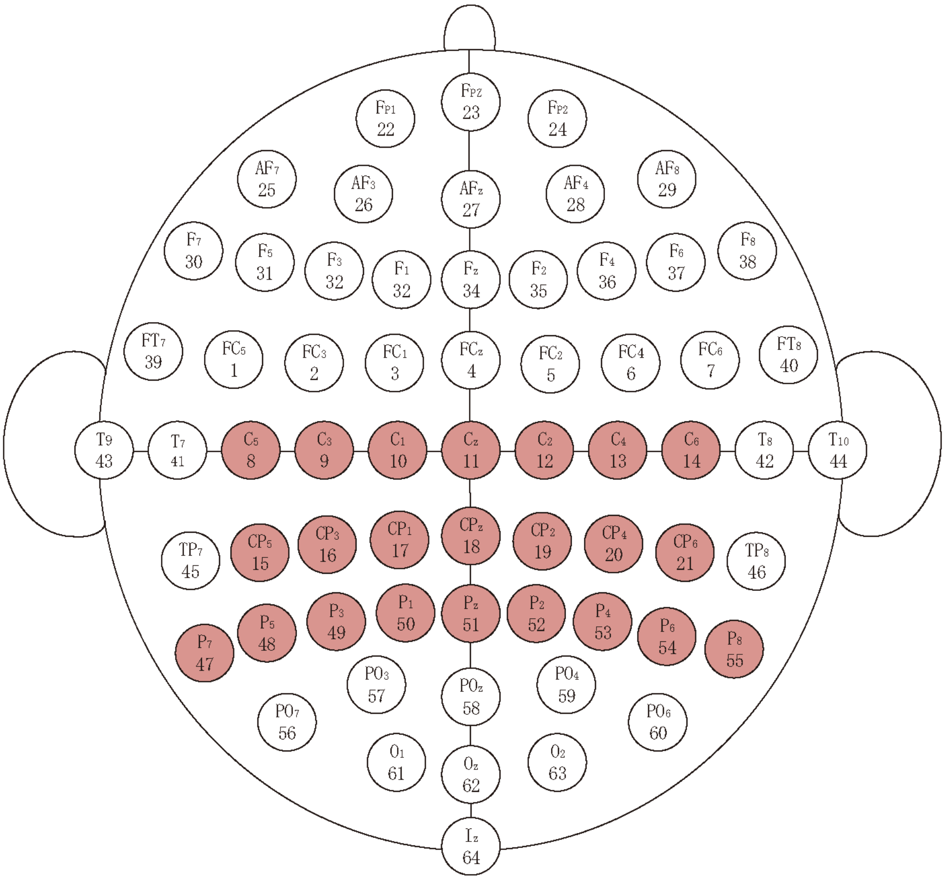

Fig. 5 *Electrode screening chart.*

C regions. However, the precision of the three models in each region varies significantly, demonstrating that the P region is capable of accurately identifying motor imagery EEG data. Compared to the single 1DCNN and LSTM model, 1DCNN-ALSTM achieves the maximum accuracy in each region, although the recognition rate gradually decreases to less than 90 % for a few electrodes. As a result, this article can utilize the 1DCNN-ALSTM model to reduce the number of electrodes, lower the cost of data collection, and increase practicability.

	Region			
Models	C [%]	CP [%]	P [%]	C-P-CP [%]
1DCNN	51.17	88.29	97.97	98.74
LSTM	29.17	28.10	30.81	80.95
1DCNN-ALSTM	94.42	94.45	98.50	99.71

Tab. II *Authentication results under different regions.*

3.3 Comparison with traditional classification networks

The output of the last fully connected layer of the convolutional neural network is usually processed by the softmax function, and the output processed by softmax is brought into the cross-entropy loss function, which can measure the degree of difference between two different probability distributions in the same random variable, and is expressed in machine learning as the difference between the true probability distribution and the predicted the difference between the true probability distribution and the predicted probability distribution. The smaller the value of cross-entropy, the better the model prediction. The cross-entropy function is essentially a log-likelihood function that can be used in both dichotomous and multiclassification tasks. In multi-classification tasks, the cross-entropy is usually of the form of Eq. (15).

$$L = -\frac{1}{n}\sum y \ln a, \tag{15}$$

where y represents the actual label, a represents the predicted output, and n represents the total number of samples.

This experiment is compared to the standard classification network, and for the purposes of comparison, the classification network's network structure is essentially identical to that of the present model. The distinction is that the classification network employs the widely used cross-entropy loss function as the network's loss function, converts the final layer of this model from a normalization layer to an output layer containing 109 neurons, and achieves classification of the input data using a softmax function. The classification network was trained using the same dataset as the current model, and once the models converged, the two models were tested using the dataset's test set.

	Indicator			
Loss function	CRR [%]	Loss	L-time [s]	T-time [s]
Cross-entropy loss	61.06	1.6061	0.3135	0.0313
Triplet loss	99.97	0.0230	0.3553	0.2369

Tab. III *Classification results of the model under different loss functions.*

As shown in Tab. III, the same 1DCNN-ALSTM model significantly improves its recognition rate when the triplet loss function is used, and the model speed is effectively increased at the same epoch, even though the model testing time is reduced. However, the reduction is within the range of engineering. In general, the triplet loss model outperforms the classic classification network in terms of identification accuracy, and the training objective is to obtain features incorporating EEG identity information, which can significantly enhance the scalability of EEG identity authentication.

3.4 Comparison with related works

The proposed 1DCNN-ALSTM-based EEG authentication model is compared to existing EEG identity models in this paper, as shown in Tab. IV. Banee Bandana

Das et al. [9] proposed a network for the resting-state only Physionet EEG motor imagery dataset in 2019, using 64 channels and 4 s of EEG for identification, with 95 % and 95.33 % for two resting states with eyes open and closed, respectively, but the real identification rate is insufficient for practical applications. In 2019, Yingnan Sun et al. [12] proposed a user identification system for EEG signals using Physionet EEG that is built on a one-dimensional convolutional long short-term memory neural network. The method's performance was validated using the motor imagery dataset, and the results indicated that the accuracy could reach 95.58 % at 16 channels. However, despite the fact that the number of electrodes has been effectively reduced, the application in reengineering remains inconvenient. In 2021, Mari Ganesh Kumar et al. [30] proposed the best subspace approach (ix-vector) to identify persons with an accuracy of 86.4 % with only 9 EEG channels. In 2022, Chiqin Lai [31]proposed a CNN combined with an error correcting output code of support vector machine with majority voting set (CNN-ECOC-SVM) for biometric recognition showing that 98.49 % accuracy was achieved in the proposed architecture. The results demonstrate that the suggested 1DCNN-ALSTM model is effective at extracting EEG authentication features and can be decreased for EEG channels, showing that the proposed method is effective, resilient, and practicable.

	Model	Electrodes	Subjecs	Time [s]	CRR [%]
Das B.B. [9]	CNN-LSTM	64	109	4	95.33
SUN Y. [12]	CNN-LSTM	64	109	1	99.58
Kumar M. [30]	IX-VECTOR	9	30	15	86.40
Lai C. [31]	CNN-ECOC-SVM	64	109	1	98.49
Proposed	1DCNN-ALSTM	64	109	1	99.97

Tab. IV *Comparison with some state-of-the-art EEG-based authentication model.*

4. Conclusions and future work

In this paper, a neural network EEG authentication model based on CNN and LSTM with attention mechanism is built, and an authentication method based on motor imagery EEG signals is designed by drawing on the triplet loss function in the field of face recognition. The performance of the proposed model is experimentally evaluated on publicly available EEG datasets, and the experimental results show that the proposed method outperforms the state-of-the-art EEG authentication methods on both 64 electrodes and 23 electrodes, thus enabling the use of less The electrodes are used to complete the authentication, which can reduce the data cost in the authentication. Furthermore, compared to traditional classification networks, the number of identifiable species in this model is not limited by the number of neurons in the output layer of the network model, which can effectively improve the scalability of EEG authentication.

In addition, EEG signals collected over long periods of time can hinder authentication performance, and thus there is still research to be done on how long it takes to update the EEG identity information. Since this paper is based on the motor imagery EEG data set for EEG authentication research, the generality of EEG can be improved in the future by implementing cross-task EEG authentication based on different tasks.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgement

The authors would like to acknowledge the authors of the dataset from PhysioNet used in this research paper and the anonymous reviewers for valuable comments.

References

[1] BERANEK B. Voice biometrics: Success stories, success factors and what's next. *Biometric Technology Today*. 2013, pp. 9–11, doi: 10.1016/S0969-4765(13)70128-0.

[2] RAMON M., VIZIOLI L. Neural microgenesis of personally familiar face recognition. *Proceedings of the National Academy of Sciences*. 2015, 112(35), doi: 10.1073/pnas.1414929112.

[3] YOON S., JAIN A.K. Longitudinal study of fingerprint recognition. *Proceedings of the National Academy of Sciences*. 2015, doi: 10.1073/pnas.1410272112.

[4] ARMSTRONG B.C., MARIA V., NEGINKHALIFIA K.J., ZHAN P.J. Brainprint: Assessing the uniqueness, collectability, and permanence of a novel method for ERP biometrics. *Neurocomputing*. 2015, 166(oct.20, pp. 59–67, doi: 10.1016/j.neucom.2015.04.025.

[5] ALEJO J., MARTEN F., MARK P., JOHN R. Characterising the dynamics of EEG waveforms as the path through parameter space of a neural mass model: Application to epilepsy seizure evolution. *Neuroimage*. 2012, 59(3), pp. 2374–2392, doi: 10.1016/j.neuroimage.2011.08.111.

[6] WANG L.Y., KONG W.Z., ZHANG X.Y. A review of brainprint recognition research. *Chinese Journal of Biomedical Engineering*. 2017, 36(5), p. 6, doi: 10.3969/j.issn.0258-8021.2017.05.013.

[7] BIEL L., PETTERSSON O., PHILIPSON L., WIDE P. ECG analysis: a new approach in human identification. *Instrumentation & Measurement IEEE Transactions*. 2001, 50(3), pp. 808–812, doi: 10.1109/19.930458.

[8] MA L., MINETT J.W., BLU T., WANG W. Resting State EEG-based biometrics for individual identification using convolutional neural networks *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2015, pp. 2848–2851, doi: 10.1109/EMBC.2015.7318985.

[9] DAS B.B., KUMAR P., KAR D., RAM S.K., BABU K.S., MOHAPATRA R.K. A spatio-temporal model for EEG-based person identification. *Multimedia Tools and Applications*. 2019, 78(2), doi: 10.1007/s11042-019-07905-6.

[10] EL-FIQI H., WANG M., SALIMI N., KASMARIK K., BARLOW M., ABBASS H. Convolution Neural Networks for Person Identification and Verification Using Steady State Visual Evoked Potential. *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2018, pp. 1062–1069, doi: 10.1109/SMC.2018.00188.

[11] ZHAO H., CHEN Y., PEI W., PEI W., CHEN H., WANG Y. Towards Online Applications of EEG Biometrics Using Visual Evoked Potentials. *Expert Systems with Applications*. 2021, p. 114961, doi: 10.1016/j.eswa.2021.114961.

[12] SUN Y., LO P.W., LO B. EEG-based user identification system using 1D-convolutional long short-term memory neural networks. *Expert Systems with Applications*. 2019, 125(JUL.), pp. 259–267, doi: 10.1016/j.eswa.2019.01.080.

[13] SCHETININ V., JAKAITE L., NYAH N., NOVAKOVIC D., KRZANOWSKI W. Feature Extraction with GMDH-Type Neural Networks for EEG-Based Person Identification. *International Journal of Neural Systems*. 2018, 1750064, doi: 10.1142/S0129065717500642.

[14] GUANG J.A. Application of unsupervised feature learning methods in EEG identity recognition. *Journal of Central South University for Nationalities: Natural Science Edition*. 2014, 33(4), p. 6, doi: 10.3969/j.issn.1672-4321.2014.04.020.

[15] MAIORANA E. Learning Deep Features for Task-Independent EEG-based Biometric Verification. *Pattern Recognition Letters*. 2021, 143(1), doi: 10.1016/j.patrec.2021.01.004.

[16] ALYASSERI Z.A.A., ALOMARI O.A., MAKHADMEH S.N., MIRJALILI S. EEG Channel Selection for Person Identification Using Binary Grey Wolf Optimizer. *IEEE Access*. 2022, 10, pp. 10500–10513, doi: 10.1109/ACCESS.2021.3135805.

[17] BENOMAR M., CAO S., VISHWANATH M., VO K., CAO H. Investigation of EEG-Based Biometric Identification Using State-of-the-Art Neural Architectures on a Real-Time Raspberry Pi-Based System. *Sensors*. 2022, 22(23), p. 9547, doi: 10.3390/s22239547.

[18] ZHANG W.Z. Study on Classification of Malware Traffic Based on Triplet Network. *Harbin Institute of Technology*. 2021, doi: 10.27060/d.cnki.ghbcu.2021.001380.

[19] WANG X.Z. CNN-based identity recognition of steady-state visually evoked EEG signals *Xidian University*. 2020, doi: 10.27389/d.cnki.gxadu.2020.002547.

[20] GOLDBERGER A.L., AMARAL L., GLASS L. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000, 101(23), p. E215, doi: 10.1161/01.CIR.101.23.e215.

[21] SCHALK G., MCFARLAND D.J., HINTERBEGER T., BIRBAUMER N. BCI2000: a general-purpose Brain-Computer Interface (BCI) system. *IEEE Transactions on Biomedical Engineering*. 2004, 51(6), pp. 1034–1043, doi: 10.1109/TBME.2004.827072.

[22] WANG S.H., PHILLIPS P., SUI Y. Classification of Alzheimer's Disease Based on Eight-Layer Convolutional Neural Network with Leaky Rectified Linear Unit and Max Pooling. *Journal of Medical Systems*. 2018, 42(5), pp. 85, doi: 10.1007/s10916-018-0932-7.

[23] SHI L.M., YANG C.Z., WU H.C. Radar signal recognition based on deep residual network and triplet loss. *Systems Engineering and Electronics Technology*. 2020, 42(11), p. 7, doi: 10.3969/j.issn.1001-506X.2020.11.12.

[24] HOCHREITER S., SCHMIDHUBER J. Long Short-Term Memory. *Neural Computation*. 1997, 9(8), pp. 1735–1780, doi: 10.1162/neco.1997.9.8.1735.

[25] SUNDERMEYER M., R. SCHLUTER, NEY H. LSTM Neural Networks for Language Modeling. *Interspeech*. 2012, doi: 10.21437/Interspeech.2012-65.

[26] GUO M.H., XU T.X., LIU J.J. Attention mechanisms in computer vision: A survey. *Computational Visual Media*. 2022, pp. 1–38, doi: 10.1007/s41095-022-0271-y.

[27] LI Q.Y., WANG P., WANG J., ZHANG T., WANG X.Y. Automatic sleep staging model based on the bi-directional LSTM convolutional network and attention mechanism. *CAAI Transactions on Intelligent Systems*. 2022, 17(3), pp. 523–530, doi: 10.11992/tis.202103013.

[28] SCHROFF F., KALENICHENKO D., PHILBIN J. FaceNet: A Unified Embedding for Face Recognition and Clustering. *IEEE*. 2015, doi: 10.1109/CVPR.2015.7298682.

[29] HERMANS A., BEYER L., LEIBE B. In Defense of the Triplet Loss for Person Re-Identification. *arXiv.org*. 2017, doi: 10.48550/arXiv.1703.07737.

[30] KUMAR M.G., NARAYANAN S., SUR M., MURTHY H.A. Evidence of Task-Independent Person-Specific Signatures in EEG Using Subspace Techniques. *IEEE Transactions on Information Forensics and Security*. 2021, PP(99), pp. 1–1, doi: 10.1109/TIFS.2021.3067998.

[31] LAI C.Q., IBRAHIM H., ABDULLAH M.Z., SUANDI S.A. EEG-Based Biometric Close-Set Identification Using CNN-ECOC-SVM. *International Conference on Artificial Intelligence and Big Data in Digital Era*. 2022, pp. 723–732, doi: 10.1007/978-3-030-97610-1_57.