Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 174 (2020) 491–504

Procedia Computer Science

www.elsevier.com/locate/procedia

2019 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2019)

EEG-based emotion classification based on Bidirectional Long Short-Term Memory Network

Jinru Yang[a], Xiaofan Huang[a], Hongkai Wu[a], Xingtong Yang[a]

Nankai University , Tianjin China
706415254@qq.com
Southwest Jiaotong University , Chengdu Sichuan Province China
1054579257@qq.com
Northeast Forestry University, Harbin, Heilongjiang, China
1251031671@qq.com
Virginia Tech, blacksburg VA,US
850194046@qq.com

These authors are contributed equally to this work

Abstract

Emotion recognition can be achieved by speech recognition, the judgment of limb movements, analysis of Electrooculogram (EOG) or capturing of facial expressions. However, those types of emotion recognition methods cannot detect human emotion well, because humankind can use fake body movement and words to hide real emotions. In this paper, we proposed an EEG-based emotion classification method based on Bidirectional Long Short-Term Memory Network (BiLSTM). Electroencephalogram (EEG) signal can detect human emotion correctly because human represent their real emotions in their mind and cannot hide emotions there. Meanwhile, EEG is a time sequence signal which needs a model which can deal with this type of data. Therefore, we chose Long Short-term Memory Network to process the EEG signal. In particular, we used an improvement version of LSTM model BiLSTM to manage the signals. BiLSTM can processes input data from front to back and back to front. Meanwhile, BiLSTM can store important information and forget unnecessary information; therefore, this process increases the accuracy of the model. Our method classifies four discrete classifications (happy, sad, fear, and neutral) for emotion classification, which achieves competitive performance compared with other conventional emotion classification methods. The final experimental results show that we can achieve an accuracy of 84.21% for four emotional states classification by using our method.

Keywords: EOG, EEG, LSTM, BiLSTM

1877-0509 © 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the 2019 International Conference on Identification, Information and Knowledge in the Internet of Things.
10.1016/j.procs.2020.06.117

1. INTRODUCTION

Emotions are ubiquitous in human's daily work and life, which is a psychological and physical state that comes from a combination of feelings, thoughts, and behaviors [1]. In recent years, the research on improving human-machine interaction system based on emotion recognition is prevalent [2]. Human-computer interaction has always been a hot topic in the interdisciplinary research field. Because human emotions are complex and mysterious, there are always many researchers working in this area. Moreover， emotion recognition has a broad prospect and application value. Emotion is a complex psychophysiological process, which can be affected by plenty of factors like the environment [3]. It can be recognized through facial expression, voice, heart rate, behavior, text, and physiological signals. Researchers can judge users' emotional state through the above contents. However, facial expressions and voice are not feasible. People may disguise themselves, sending out inaccurate emotional signals to hide their real feelings. However, the emotion recognition of physiological signals, such as the detection and analysis of human's emotions through identifying brain waves, can avoid misleading by facial expression well and make emotional recognition more effective.

In EEG analysis, traditional methods include wavelet analysis [4], matching and tracking, chaos analysis [5], but these methods need to design and extract the characteristics of EEG signals manually. Those methods have high redundancy and fail to take into account the temporal dynamics of EEG signals, which are crucial for emotion recognition. Long Short-Term Memory (LSTM) network adopted in this paper is a revision of the Recurrent Neural Network (RNN). It can overcome the weakness of the traditional deep learning algorithm in dealing with temporal information.

BILSTM is a two-way LSTM, that is, a combination of forwarding LSTM and backward LSTM. The LSTM captures the connection between the front and back of the information better [6].

Shanghai Jiaotong University did an experiment which contained 15 participants to watch 24 groups of videos about 2 minutes each which have the tendency of inducing happy, sad, fear, or neutral emotion [7]. While watching those videos, participants' EEG signals and eye movements are collected with the 62-channel ESI NeuroScan System. The original EEG data were filtered out of noise and artifacts, and the differential entropy (DE) characteristics were extracted in 5 frequency bands. According to this experiment, they create a dataset called SEEDIV to note the experimental data.

In this paper, we focus on the emotional classification of analyzing temporal EEG data using LSTM. After the acquisition of EEG data and extracting DE features, the Linear dynamic system (LDS) and moving average methods were used to process the data. In the end, we apply LSTM for emotional classification.

2. RELATED WORK

2.1 Emotion Recognition based on EEG

People's research on emotion recognition mainly divides into two categories: recognition based on non-physiological signals [8] and recognition based on physiological signals [9]. Recognition of emotions based on non-physiological signals mainly includes facial expressions, body language, and voice intonation.[10] For example, the smell means happy, and tears mean sad. Yu-Dong Zhang [11] used biorthogonal wavelet entropy to extract multiscale features and applied fuzzy multiclass support vector machine as the classifier to recognize emotion based on facial expression images.

Researchers emotion can be recognized by audio. For example, when people feel happy, they will speak much faster than their average speed; When people feel angry, they will speak louder than usual voice [10]. Mirsamadi S et al. [12] employed RNN to discover emotional relevant features from speech automatically.Since EEG signals are objective and it is difficult for people to disguise their emotions in EEG signals, this kind of method is much accurate than non-physiological signals.

Some researchers choose to use EEG signals to classify emotions. For example, the study of Naiyu Wu and others [13] shows that EEG signals can make the classification of emotions more effective.

Researchers have applied multiple methods on EEG signals to analyze emotions. Wu Naidu et al. [13] used PCA based emotional feature extraction method and SVM based emotional feature matching algorithm to classify emotions. Emotions were divided into positive and negative categories. Cheng et al. [14] applied typical spatial pattern (CSP) algorithm based on emotional characteristics of vocabulary materials to examine the differences between depression and universal emotions according to brain waves induced by Chinese words. Relaxation et al. [15] used fMRI assisted EEG channel selection method to study emotional classification.

2.2 Data processing method：Deep Learning

Deep learning is a research direction in the field of machine learning and the most popular research topic in the field of machine learning [16]. The concept of deep learning comes from the study of how the human brain learns knowledge [16]. By simulating the multi-layer neurons of the human brain to abstract concepts and interpret the mechanism of data, neural networks are established to realize the automatic feature extraction of input data [17]. The advantage of deep learning is that by layer-by-layer extraction, lower-level features can be abstracted into more expressive high-level features, and data features are extracted through hidden layer nodes. Therefore, deep learning is widely used in data analysis in various fields. It has been a great success in image classification, video analysis, computer vision [18], speech recognition[19], natural language processing[20], and many other fields. That is why our experiment chose to use deep learning to deal with EEG-based emotion classification.

What distinguishes deep learning evidently from traditional pattern recognition is the application of automatic learning based on big data. The special characteristics make it possible that the outcome with pattern recognition approach would be promoted [21]. Over the past few decades, the characteristics of manual design have been dominant in a variety of applications for pattern recognition [22]. Because manual design relies primarily on the designer's prior knowledge and the characteristics of big data, big data does not take advantage of itself. The number of parameters allowed to appear in the design of features is minimal because of the reliance on manual tuning parameters. Almost all deep learning models can automatically learn the representation of features from big data and deep neural network can contain thousands of parameters. The model we chose this time is BILSTM, which will be introduced in the next section. There is an apparent contrast between human learning and machine learning: manual design often takes five to ten years to design useful features by hand [23]. However, in the framework of neural networks, BILSTM can quickly learn new feature representations from data within several hours. This is the main reason why we chose BILSTM to identify and classify emotions in this experiment.

2.3 Algorithm

When dealing with different subjects, different models have been developed to more precisely fit and analyze the dataset. Natural language processing is typically about speech recognition, language modeling, machine translation, etc. In this processing, the chronological order in which samples appear is critical [24]. In order to adapt to this demand, a neural network structure has emerged - the Recurrent neural network (RNN). In RNN, the output of neurons can act directly into itself at the next timestamp. By continually feeding the model data and training, RNN can solve natural language processing or computer vision and other related problems. [24]

Because the memory of RNN is actually limited, it can only memorize the information of the previous steps, and as the sequence grows, it is prone to the problem of gradient explosion and gradient disappearing [25]. The researchers created a specialized recursive neural network called Long-Short-Term memory (LSTM). LSTM is essentially like RNN, but they can use a mechanism called " gate " to learn about long-term dependency.

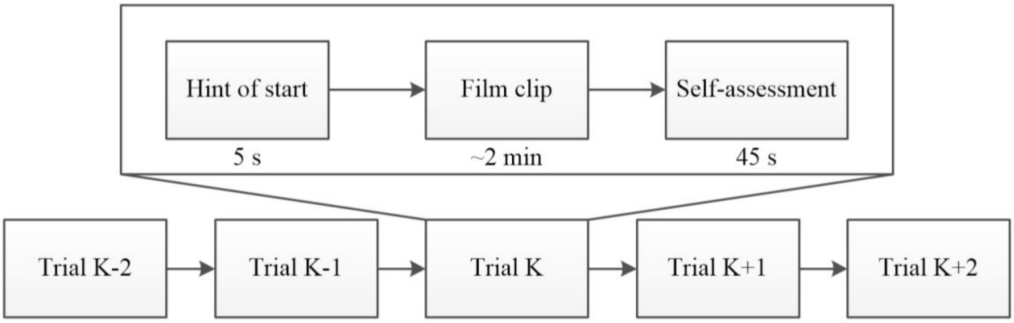
LSTM, Long Short-Term Memory is a time-cycle neural network designed to solve the long-term dependence of RNN. Just like the problem we are dealing with in this paper "emotion states classification" both time series problems and time period problems, the model can pass relevant information in the sequence processing. LSTM can learn the long-distance information in the data set by memorizing the previous cell unit information and updating the node information with the current cell state value, and can select the information we want to output by forgetting the irrelevant information [25]. Therefore, the LSTM network is a good choice for emotion classification from EEG.

The addition and removal of information is achieved through the "gate" structure, which has three types of gate structures: forgetting gates, input gates, and output gates. The LSTM algorithm can achieve the goal that we need to learn Long-Short-Term dependencies. Needless to say, there are many special models in LSTM to adapt to different requirements. As a result, the LSTM model became the first choice for our experiment.

3. METHODS

3.1 Data Preprocessing

The most crucial thing in data preprocessing is to limit the size of data. We have 15 participates, and each of them would watch 24 different pieces of two minutes videos. When the researchers of Shang Hai Jiao Tong University did this experiment, they used three weeks to record data from those 15 participates. In our experiment, we used the data from the first week to test our model [7].



Fg.1 Procedure of emotion data collecting [7]

When participates watch each video trial, they will be given 5 seconds hint at the start to tell them the emotion that video will present. Then, the participates watched a 2 minutes standard emotion's film clip. In the end, they have 45 seconds to record their emotion.

There are lots of different features can be extracted from EEG signals such as energy spectrum (ES), differential entropy (DE), rational asymmetry (RASM), and differential asymmetry (DASM) [26]. In the paper from Ruonan Duan et al. they find that DE has the best performance in average accuracy on classification processing. So, in our experiment, we only use the DE feature from the dataset [26]. After getting DE feature from the dataset, we select the Linear dynamic system (LDS) to do the smooth process.

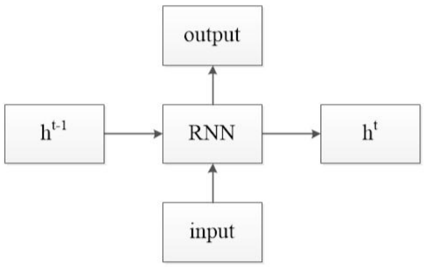
The 24 videos trials are not the same size. However, the input size of the LSTM model must be the same. In this case, we set the 24 videos into the same size by doing the zero-padding through MATLAB. We found that the most extended video is 64 time unites, so we add zeros into other metrics to ensure the coherence of the data. The reason why we choose to add zeros into the matrix is that people may represent their emotions at different time. Some people may expose their emotion at the last 10 seconds; Some people may expose their emotion right after the end of the video. In this case, using the shortest or average time of the video is not a logical choice. As a result, we chose the most extended video and zero pad other matrix to be the same size as the most extended video. The creators of the SEEDIV dataset use the model with 62 channel ESI NeuroScan System to record the EEG signals [7] and each signal separates into five frequency bands. As a result, the size of our data is a three dimensions matrix which is 64x62x5.

Each video of the 24 video sets stands for one kind of emotions. There are four types of emotions in this research, which are neutral, sad, fear, and happy. One video represents one kind of emotion. The creator of the dataset labeled these videos with their emotions by numbers, for example, 0 is neutral, 1 is sad, 2 is fear, and 3 is happy. Therefore, when we input our dataset to LSTM model, it should have been correctly labeled. So, we add the 360 labels and our old three dimensions matrix together, then created a new data set 360X64X310. After we confine the data into the same size, we can send the samples into our LSTM model.

Considering that we collect EEG from 15 people, and different people or even the same person at a different time may response diversely to the same video, the EEG to the same video would have different intensity and patterns. Therefore, we choose to apply standardization on the data set to scale it into a certain realm, in order to avoid the influence of extreme data and accelerate the speed of learning.

3.2 Bidirectional LSTM Networks

Conventional neural network like CNN does well in extracting invariant features. However, when it comes to predicting the current output conditioned on long distance features, RNN performs better than CNN[27]. RNN is good at modeling units in sequence because it maintains a memory based on history information. In every RNN unit, the input represents features at time t and has the dimensionality as feature size(Fg.2). Hidden state h^{t-1} represents memory before this unit. Using h^{t-1} and input, we can calculate and pass new memory h^{t} to the next RNN unit. RNN also has its own weakness,which just mechanically calculates the sequential information one by one, regardless of different level of influence of those information. Therefore RNN has difficulty in finding and exploiting long-term dependences in the data set [28]. When dealing with sequential information, RNN sometimes meets with problems like gradient explosion and gradient vanishment.



Fg.2 Structure of RNN Cell

LSTM is a method developed from RNN, which has a better performance when dealing with long-term memory. Different from RNN, the hidden layer of LSTM updates by a new memory cell(Fg.3). As is shown in the figure, i means input gate, o means output gate, f means forget gate and c means cell vector. The LSTM memory cell is calculated by the followings:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (1)$$

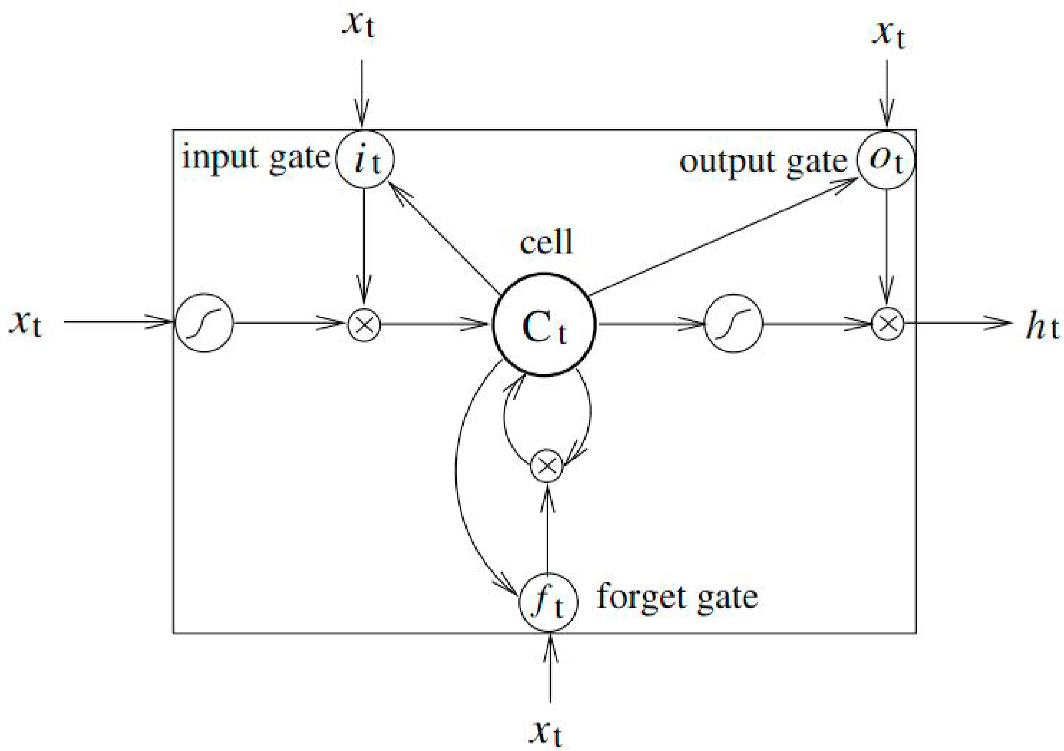
$$i_t = \sigma\left(W_{x_i}x_t + W_{h_i}h_{t-1} + W_{c_i}c_{t-1} + b_i\right) \quad (2)$$

$$f_t = \sigma\left(W_{x_f}x_t + W_{h_f}h_{t-1} + W_{c_f}c_{t-1} + b_f\right) \quad (3)$$

$$c_t = f_t c_{t-1} + i_t \tanh\left(W_{x_c}x_t + W_{h_c}h_{t-1} + b_c\right) \quad (4)$$

$$O_t = \sigma\left(W_{x_o}x_t + W_{h_o}h_{t-1} + W_{c_o}c_t + b_o\right) \quad (5)$$

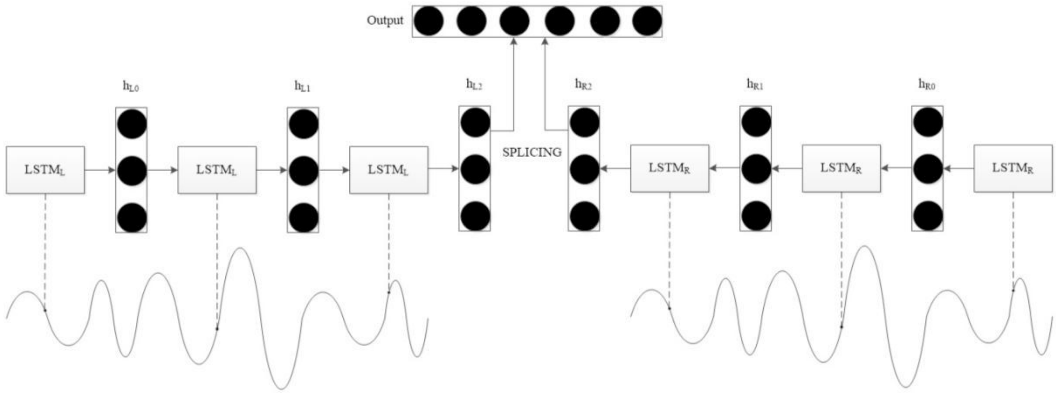
$$h_t = o_t \tanh\left(c_t\right) \quad (6)$$



Fg.3 LSTM Cell [28]

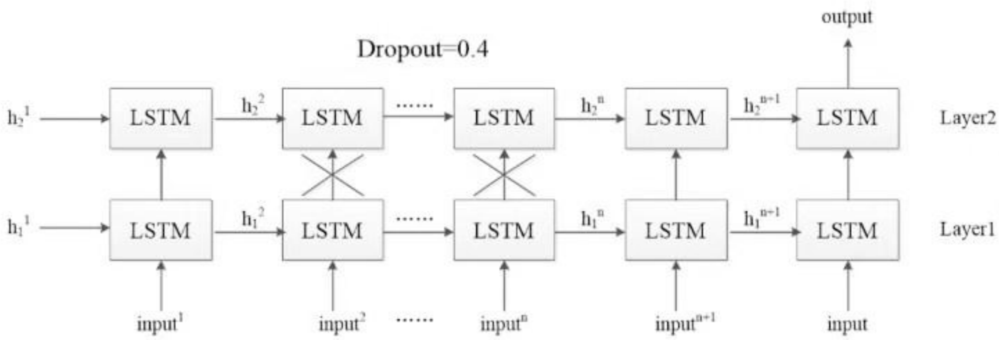
The matrixes in the above formulas have the meaning as their names suggest. For example, W_{h_o} means hidden-input gate matrix. After the processing in LSTM unit, it passes hidden state h^t to the next LSTM unit which is responsible to solve the signal of next time slice and its output to the next layer. Because LSTM maintains some special units like forget gate, this method can not only remember important long-term memory but also be adapted to short-term memory which contains important information.

In our research, we applied BiLSTM to construct our model, because some research has revealed that the latter EEG contains some information about former EEG [29], but conventional LSTM can only handle unidirectional timing signal. BiLSTM is a revision of LSTM, (Fg.4) it sends not only the positive sequence of EEG but also the inverted sequence of EEG to the learning method, works just like combining two LSTM together.



Fg.4 BiLSTM Network

In our experiment, we set 64*2 LSTM units because we scaled data set into 64 time unites, every second has one information, and BiLSTM doubled the size of inputs and outputs. We set input layer size as (64,310) because in every time step the feature vector is 310. We set the hidden layer size of BiLSTM as 1024 to balance time-consuming and extracting enough features. As we can see from Fg.5, when dropout is 0, every LSTM cell in one layer will pass its output to the next layer. Dropout ranges from 0 to 1. The higher dropout is, the less LSTM cell will pass its output to the next layer. In order to prevent overfitting and decrease the complexity of the model, we set dropout as 0.4(Fg.5).



Fg.5 How Dropout Works

3.3 Dimensionality Reduction and Classification

The output feature dimension of BiLSTM is 2048 features (1024 from the positive sequence and another 1024 from the inverted sequence). Each sample corresponds to one certain emotion of four kinds, therefore we must reduce dimension of features from 2048 to 4. During our experiment, we found that when we set 3 linear transformation layer as 3, our model will extract the most information and perform best in testing.

Linear transformation:

$$y = xA^T + b \tag{7}$$

In order to classify emotion state every sample belongs to, we apply softmax to convert 4 values we get from linear transformation to four probability the sample belongs to each kind of emotion. The highest probability represents the given emotion for the trail EEG data. Then we apply cross-entropy to calculate the loss and adjust the parameters of our model. The loss points out which direction the parameters are adjusted to and learning rate decides the length it goes. After 150 epochs our model gradually grows stable and can be applied to testing.

Softmax:

$$\text{Softmax}(x_i) = \frac{exp(x_i)}{\sum_{j=1}^{K} exp(x_j)} \tag{8}$$

4. EXPERIMENTAL RESULTS

In this paper, we use SEEDIV data set to verify the learning efficiency of our BiLSTM model. We selected DE features from SEEDIV data set for later analyzing because DE features have higher accuracy and lower standard deviation than other features, which means that DE features are more suitable for emotion classification based on EEG.

When using the DE characteristics of the total frequency band as data set, the weighted accuracy of emotion recognition of BiLSTM model is 84.21%. We choose weighted accuracy to evaluate learning efficiency because the weighted average accuracy can better reflect the classification ability.

We compared our model with two other models which also use SEEDIV as data set. Zheng W L et al. [30] applied three conventional pattern classifiers, namely k nearest neighbors(KNN), logistic regression (LR), and support vector machine (SVM) and a newly developed pattern classifier, discriminative Graph regularized Extreme Learning Machine (GELM) separately to divide emotions into 4 consecutive categories, with an average accuracy of 79.28%, which was lower than our average accuracy. Zheng W L, Member S et al. [7] applied multimodal emotion recognition framework to divide emotions into 4 discrete categories, and the average accuracy data of emotional classification based on EEG was 70.33%. Our model performs better than both of the methods above(Table.1).

Table.1 Accuracy comparison

Methods	Average Accuracy
Zheng[30]	79.28%
Zheng[7]	70.33%
Our	84.21%

To sum up, combined with the results of previous similar studies, the learning efficiency of our BiLSTM model is much higher than that of previous studies. The detailed testing results are as follow (Table.2).

Table.2 Experiment result

Emotion	precision	Recall	F1-score	support
Neutral	0.83333	0.8333	0.8333	12
Sad	0.76667	1.0000	0.8679	23
Fear	0.81818	0.8182	0.8182	22
Happy	1.00000	0.5333	0.6957	15

Recall is the ratio of the number of samples retrieved correctly to the number of samples that should be retrieved.
Recall:

$$\text{Recall} = \frac{TP}{(TP+FN)} \tag{9}$$

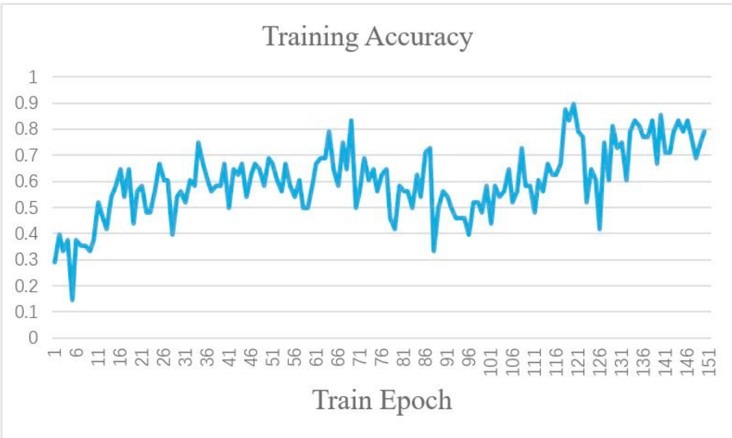
Among the four categories of emotions, Sad has a recall of 1, which means that the retrieval is completely correct, while Happy has a recall of 0.5333, the lowest, which means that the retrieval is only half correct.
F-score is a comprehensive evaluation index, which refers to both Precision and Recall.
F-score:

$$\text{FS} = (1+\beta^2)\cdot\frac{Precision{\cdot}Recall}{\beta^2{\cdot}(Precision+Recall)} \tag{10}$$

In the formula, β is used to adjust the weight. When $\beta = 1$, weights of Precision and Recall are the same and the formula is abbreviated as $\text{F1} - \text{score}$. The $\text{F1} - \text{score}$ of Sad, 0.8679, is the highest; The $\text{F1} - \text{score}$ of happy, 0.6957, was the lowest.

In addition, we analyze the stability of the model during operation.

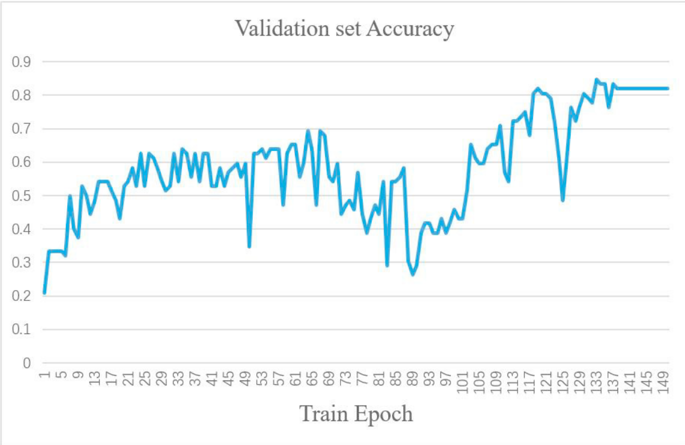
As is shown in Fg.6, with the increase of the train epoch, training set accuracy gradually increases. As is shown in Fg.7, with the increase of the train epoch, training set loss gradually decreases, indicating that the BiLSTM model gradually converges in training with good stability.



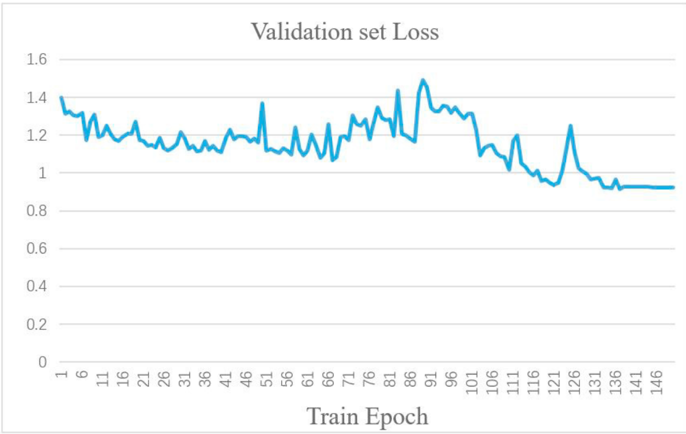
Fg.6 Training Accuracy

Fg.7 Training Loss

Shown in Fg.8, with the increase of the "train epoch", the accuracy of the validation set gradually increases. As is shown in Fg.9, with the increase of the "train epoch", the validation set loss gradually decreases, indicating that the BiLSTM model of the validation set gradually converges and has good stability.



Fg.8 Validation set Accuracy



Fg.9 Validation set Loss

In conclusion, the BiLSTM model gradually converges in both the training and the validation set, with good stability.

5. CONCLUSION

In our research, we use the EEG signal to recognize human emotion other than using facial expression. EEG is a time sequence signal, which needs a model which can process time sequence signal. To pursue a higher accuracy, LSTM stands for the right choice for dealing with EEG signal. LSTM can increase the accuracy of classification because LSTM has a system called cell. The cell contains four gates to help us filter the unnecessary information and keep critical information along with input data. Meanwhile, we use BiLSTM to boost accuracy and efficiency for our final results. BiLSTM is an improved model based on LSTM, which has the capability to process input data twice from different directions and increase training efficiency.

As a result, the most identical method in our research is that we use a model (BiLSTM) which can handle the time sequence signal (EGG) of human emotions to detect four discrete human emotions: happy, sad, fear, and neutral. After applying all the data into our model, we got 84.21% weighted accuracy on clarifying four emotion at the final simulation.

Additionally, we can extract data of eye feature and combine the eye feature data with EGG feature data. In this way, we can get a much precise result than only using EGG feature data. Moreover, we put all the five bands of DE together to train our model. Some research has shown that different bands may represent different kinds of emotion information. If we separate them, train the models according to them separately and combine the models in the end, that new model may perform better than our present one. Last but not least, we can also refine our learning model. For instance, some revisions of LSTM like BiLSTM-CRF may perform better than conventional LSTM, and we can also define our loss function and validation to make the learning method more suitable to our task, emotion classification.

Acknowledgements

The completion of the thesis is attributed to many people's support and encouragement. Without of the help of many people,it could never be completed, thus here we would like to express my sincere gratitude towards them.
First and foremost, we owe our heartfelt thanks to my distinguished and cordial supervisor,Professor Cuixia Ma, who influenced us with his insightful ideas and meaningful inspirations, guided us with practical academic advice and feasible instructions, and enlightened us while we was confused during the writing procedure. His thought-provoking comments and patiently encouragements are indispensable for my accomplishment of this thesis.
Ultimately, thanks go to our parents who are our mentor and guardian from the very beginning in primary school. Without their refined education and care,we could never grow up in such a joyous and cozy environment nor have the courage to confront any obstacles on our way to success.

References

[1]Cornelius R R . The science of emotion: Research and tradition in the psychology of emotion[J]. PsycCRITIQUES, 1996, 42(3).

[2] Andreasson R , Alenljung B , Billing E , et al. Affective Touch in Human−Robot Interaction: Conveying Emotion to the Nao Robot[J]. International Journal of Social Robotics, 2018, 10(4):473-491.

[3]Storman, storman, wang l. emotional psychology: from daily life to theory [M]. China light industry press, 2006.

[4] Sai C Y , Mokhtar N , Arof H , et al. Automated Classification and Removal of EEG Artifacts with SVM and Wavelet-ICA[J]. IEEE Journal of Biomedical and Health Informatics, 2017, PP(99):1-1.

[5] Yang S , Zhang A , Zhang J , et al. A new chaotic feature for EEG classification based seizure diagnosis[C]// 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2017.

[6]Tao C, Xu R, He Y, et al. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN[J]. Expert Systems with Applications, 2016:S0957417416305929.

[7]Zheng W L, Liu W, Lu Y, et al. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions[J]. IEEE Transactions on Cybernetics, 2019, 49(3):1110-1122.

[8] Byoung K . A Brief Review of Facial Emotion Recognition Based on Visual Information[J]. Sensors, 2018, 18(2):401-.

[9] Lin S , Jinyan X , Mingyue Y , et al. A Review of Emotion Recognition Using Physiological Signals[J]. Sensors, 2018, 18(7):2074-.

[10] Bao xanni, he jin-cai. Effects of emotional and brain injury on color recognition ability and preference [C]// zhejiang association of psychosomatic medicine conference. 2006.

[11] Zhang Y D , Yang Z J , Lu H M , et al. Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation[J]. IEEE Access, 2017, 4(99):8375-8385.

[12] Mirsamadi S , Barsoum E , Zhang C . Automatic Speech Emotion Recognition Using Recurrent Neural Networks with Local Attention[C]// ICASSP. IEEE, 2017.

[13] Wu Naidu. Emotional classification research based on EEG signals [D]. Central University for nationalities, 2013.

[14] Cheng Minmin. Study on characteristics of emotional information based on EEG and its classification method [D]. 2017.

[15]Rresearch on emotional recognition method based on EEG [D]. PLA University of information engineering, 2015.

[16] Li Zhenqi. Research and system realization of automatic emotion recognition based on multi-domain features of EEG time-space frequency and frequency [D].

[17] Wei Jian, Liu Aijuan, Tang Jianwen. Research on Alarm Model of Digital TV Monitoring Platform Based on Deep Learning Neural Network Technology[J]. Cable TV Technology, 2017(7).

[18] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G. and Cai, J., 2015. Recent advances in convolutional neural networks. arXiv preprint arXiv:1512.07108.

[19]Xu Tingting, Ma Chenxi, Zhang Jingmin, et al. LAMOST Spectral Classification Based on Deep Learning[J]. Acta Astronomica Sinica, 2019, 60(02).

[20]Mikolov T , Chen K , Corrado G , et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.

[21] J. Deng, W. Dong, R. Socher, and et al.. Imagenet: A large-scale hierarchical image database. In IEEE Int′l Conf. Computer Vision and Pattern Recognition, 2009

[22] Yang Mei, Chen Ning. Covering song recognition model based on deep learning and hand-designed feature fusion[J]. Journal of East China University of Science and Technology (Natural Science), 2018, 44(05):138-145.

[23] Yi Zhuang. Interpretation of Research Progress and Trends of Big Data from the Perspective of Database[J]. Electronic Technology and Software Engineering, 2014(17):206-206.

[24] Goodfellow, I., Bengio, Y., Courville, A. . Deep learning (Vol. 1) : Cambridge : MIT Press , 2016 : 367-415

[25] Deng Yang. Using deep learning method for Chinese microblog sentiment analysis research. [D]

[26] Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu,″ Differential Entropy Feature for EEG-based Emotion Classification″ in 6th Annual International IEEE EMBS Conference on Neural Engineering. IEEE 2013.

[27]Yin W , Kann K , Yu M , et al. Comparative Study of CNN and RNN for Natural Language Processing[J]. 2017.

[28]Huang Z , Xu W , Yu K . Bidirectional LSTM-CRF Models for Sequence Tagging[J]. Computer Science, 2015.

[29]Fares A . Region level Bi-directional Deep Learning Framework for EEG-based Image Classification[C]// 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2018.

[30]Zheng W L , Zhu J Y , Lu B L . Identifying Stable Patterns over Time for Emotion Recognition from EEG[J]. IEEE Transactions on Affective Computing, 2017:1-1.