Inter-subject Attention Recognition from Multi-Channel EEG with Transfer Learning based CNN-LSTM model

Team: Xingqi Pan, Meiyi Yu, Haiqin Zhao**. Project Mentor TA:** Maheshwarran Karthikeyan

1) Abstract

For this project, we tried to detect the human mental states from Electroencephalogram (EEG), a most studied time-sequential signal which has been proven to have the capacity to detect attention, with a Transfer Learning based state-of-the-art deep learning model, to improve study/work efficiency for various people in different scenarios. In this research, we proposed a framework based on the CNN-LSTM model which learns the inter-subject attentive mental state from multi-channel EEG signals. We also delivered an innovative method of EEG signal preprocessing which enhances the correlations among data without adding redundant information. The result shows that the model reaches an average accuracy of 69.36% on 5 participants for leave-one-out inter-subject attention recognition on a public dataset, which is higher than the similar CNN-LSTM model of 65.29% on single-channel EEG signals. It also points out the possibility of applying multiple EEG channels on mental state detection.

2) Introduction

Attention is one of the key parameters of human consciousness which is highly related with study and work efficiency. It is used by people to control details of their awareness and is associated with people's voluntary ability to select among competing items, to correct mistakes and to moderate and adjust their emotions[1]. In the past several years, with scientific and clinical interest increasing, researchers pay attention to appling EEG signal for characterizing and quantifying the neurophysiology of attention-deficit/hyperactivity disorder (ADHD), which shows a possible connection between EEG signal and attention [2]. Electroencephalogram (EEG) is a most studied time-sequential signal of recording brain activity using electrophysiological indicators [3,4]. Recently, researchers introduced Deep Learning algorithms in the domain of EEG signal processing for human mental state recognitions. However, most existing models try to deal only with single channel EEG data, but with the scarcity of dataset, it may not always be possible to find demand signal channels.

In this research, we aim to build a subject-independent learning system, which learns EEG signals from multiple subjects to recognize the attention of the test target, to help people switch their mental states quickly between focus and unfocus. As EEG signals may always be interrupted by other human activities [5], several human activities may affect the accuracy of EEG, hence we introduce a new way of data preprocessing method to deal with the unexpected noises, basically using several filters at first to pick the useful information out and then split into time segments for experiment. To deal with multi-channel signal processing, we designed a transfer learning based attention pattern recognition model from multi-channel EEG based on the current existing CNN-LSTM network using the idea of Transfer learning. In model structure, the Convolutional Neural Network (CNN) layer, as one of the essential machine learning models, is used for feature extraction and a Long Short-Term Memory Layer is implemented for data classification [7,8].

3) Background

EEG with Machine Learning is widely used in brain-computer Interface for medical diagnoses, such as attention deficit hyperactivity disorder(ADHD), epilepsy, sleep disorders, etc. In the previous years, many studies on EEG signals have used traditional machine learning to uncover relevant information for neural classification and neuroimaging. [6] For example, independent component analysis (ICA) for artifact removal, principal component analysis and local Fisher's discriminant analysis (LFDA) for feature dimension reduction [9]; Support Vector Machine and Random Forest for EEG signal Classification.[10] Recently, researchers started to apply machine learning techniques on EEG signal analysis with the advent of deep learning. In [11], the team introduced an MLP neural network as a classifier for the analysis of early diagnosis of ADHD from a single Fp1-A1 EEG signal. This technology is also widely used in the field of mental workload measurement. [12] provides an unsupervised learning model for predicting driver drowsiness from Oz channel. [13] presents an application of DCNN on detecting the attentive mental state from Fp1-Fp2 raw EEG data. [14] delivered a state-of-the-art CNN-LSTM model for detecting subject-independent driver drowsiness, using Oz channel EEG signal collected from 27 subjects performing VR driving tasks. Also, the previous research also shows that the attention level is associated with the signal collected on the posterior cortex in the alpha, beta, theta, and delta band[9].

Our project is built based on [13], and [14]. In our research, different from previous research, we used an EEG signal dataset generated by O1 and O2 electrodes, two electrodes which are located close to Oz and can also provide information about attentive mental states, for multi-channel research. We built models for both single-channel data and multi-channel data for performance comparison.

4) Summary of Our Contributions

1. Contribution(s) in Code: N/A
2. Contribution(s) in Application: We modified the existing CNN-LSTM model and applied it in a new field. Previously, the EEG concentration prediction was mostly trained on driver's tasks, but we tailored the CNN network and studied the students' concentration status instead.
3. Contribution(s) in Data: We introduced a new data preprocessing way on EEG signal. The EEG dataset we are dealing with is multichannel signals for length of around 40 minutes. When we cleaned the data, we reserved most of the information by reusing some signal fragments instead of just doing replication padding.
4. Contribution(s) in Algorithm: N/A
5. Contribution(s) in Analysis: N/A

5) Detailed Description of Contributions

5.1 Dataset and Features

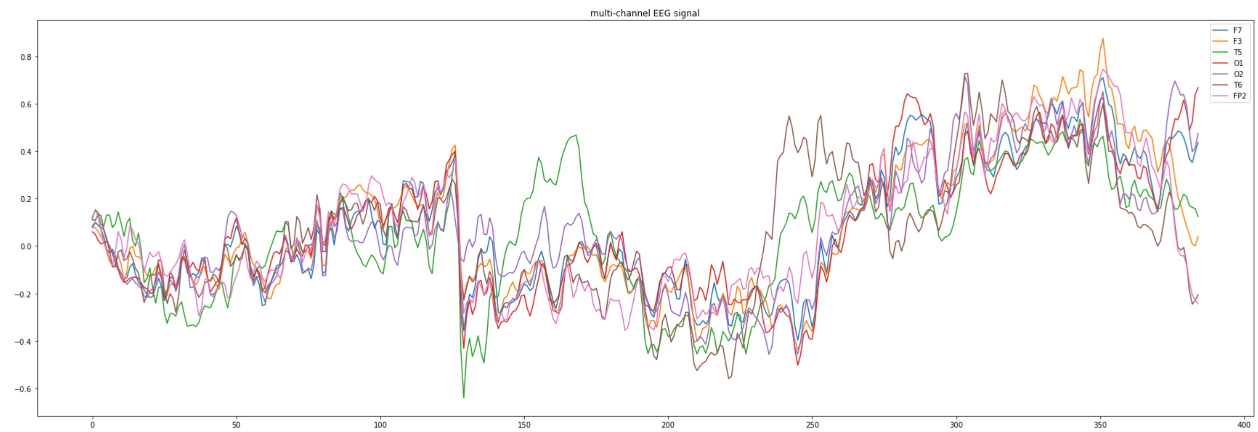

Figure 1 Multi-channel EEG signal

This study used a raw EEG dataset provided by [15] which conducted 34 experiments in 5 subjects for 7 separate days to monitor their mental status as focused, unfocused, drowsy by EMOTIV devices. The signals were collected from 34 experiments for monitoring of attention state in human individuals using passive EEG BCI. There are 14 electrodes AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 corresponding to 14 signal channels. The sampling rate of the signals is 128Hz. Each of the signals record the human attentive mental states over 20 minutes and can be splitted into 3 parts: first 10 minutes are measuring for ‘focus’ state, 10-20 minutes for ‘unfocus’ state, and the left minute as ‘‘sleep’’’ state.

5.2 Methods

5.2.1 Data Preprocessing

In our study, our goal is to differentiate people’s concentration levels. To be more specific, we want to tell whether one is focused or unfocused. So we will only utilize the first 20 minutes of the dataset. The size of the data for each experiment after processing is 398*2*447. Based on [13] and [14], Here is our data-preprocessing method:

1. To perform multi-channel analysis, we picked O1, O2 [12] electrodes as our data input, since the occipital (O) region can best represent the concentration level. We stacked these two signals together to inject into the neutral network.
2. Filter the signals to 0.5-40Hz in frequency domain. One reason is that the five classical bands δ (0.5–4 Hz), θ (4–8 Hz), α (8–12 Hz), β(12–30 Hz), and low ϒ(30-40Hz) are most associated with the changes in individuals’ attention state [15]. The other reason is that human brain signals contain many low-frequency artifacts/noise that need to be eliminated.
3. Applied a 3s sliding window on each signal channel, that is, each data has 3s*128Hz = 384 features, which preserve the data information to the greatest extent. To prevent the loss of relations between data after passing through the Convolutional layer, we expanded data to 447 features with some overlapping data sliced from previous and late segments of the continuous EEG time series. Removing the first and the last data in each channel as they are just used for segment expendation.
4. Combining all signal slices from different channels together in time domain as the multi-channel training set. Also, generating labels where focused data labeled with 0 and unfocused data labeled with 1.

5.2.2 Design of network

As EEG is a time sequence model, the data on the future stage is dependent on the previous stage. Hence, we used a CNN-LSTM model for measuring drowsiness [14] as the prototype. This model combined both CNN for compress data and LSTM for feature extraction which get an average accuracy of 72.97% on the driver drowsiness set. To promote the model on the set we selected, we change its architecture to fit the multi-channel signal data and try to get better performance on subject-cross classification[16]. The state-of-art data analysis was based on single-channel EEG signals, which only represents activities of specific regions of one's brain. As an improvement, we prepared signals from two different electrodes(O1 and O2) as CNN inputs, and improved the CNN to multi-channel structure, so as to acquire more information and improve the accuracy.

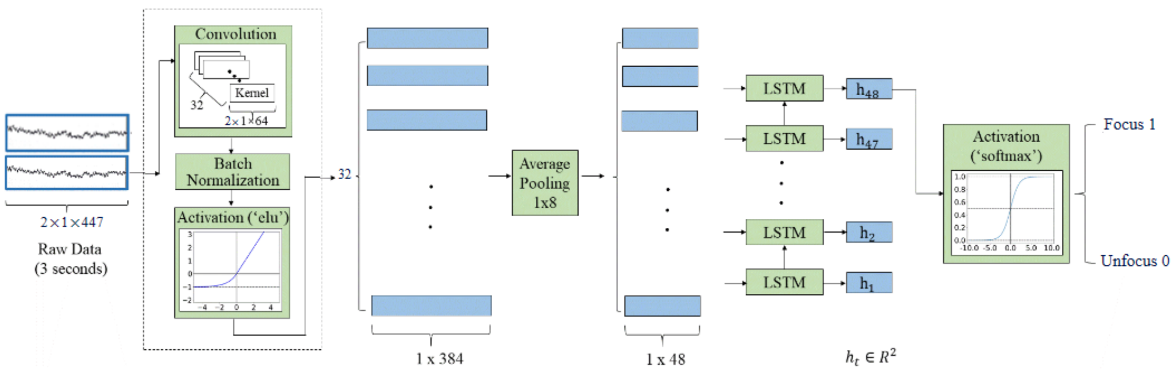

Figure 2 CNN-LSTM Architecture

- The convolutional layer contains 32 kernels with stride 1. Each 2-layer 1-dimensional kernel has the length of 64, as with half of the sampling rate(128Hz), it can capture essential EEG frequencies above 2Hz. The input data has the length of 447 contains the data in 3s period with overlapping signals from the original data to maintain the most information.
- The Batch Normalization layer can remove the internal covariate shift.
- The average pooling layer has the window size of 1 * 8, it can summarize the extracted information and prevent overfitting.
- The LSTM layer maps the feature sequence to hidden states to learn the temporal dependency of the feature sequence.
- The final activation layer performs Softmax function on hidden states to output the classification results.
- The Adam optimizer has been selected for minimizing the loss function properly.

To measure the accuracy on inter-subject level, for each independent experiment, we used a approach call Leave-one-experiment-out cross validation, because in this way we can better showcase the model's capabilities of cross-subject classification. Under this approach, each time the data from one experiment is used for testing, while the data from all the other experiments are used for training. This method encourage the trained model to be able to reveal the concentration status when some new task is performed by some new person.

5.3 Experiments and Results

Different from previous research, we picked a more randomly designed EEG dataset which recorded the change of EEG signals when subjects performed study tasks. Our research goal is to verify whether the current model can also give good prediction results in this new dataset and new application field, and improve the network structure to train better performance. Before the experiment, we made a hypothesis that a multi-channel dataset contains the same or more information, compared with a single channel dataset, hence the multi-channel model has the same performance as the single-channel model. We preprocess and clean the dataset, getting a 398*34 data frames with each column as a trial, and every entry is a 3s signal slice. Two channels(O1, O2) of the original data were being preprocessed and stacked together in the time domain as our training set.

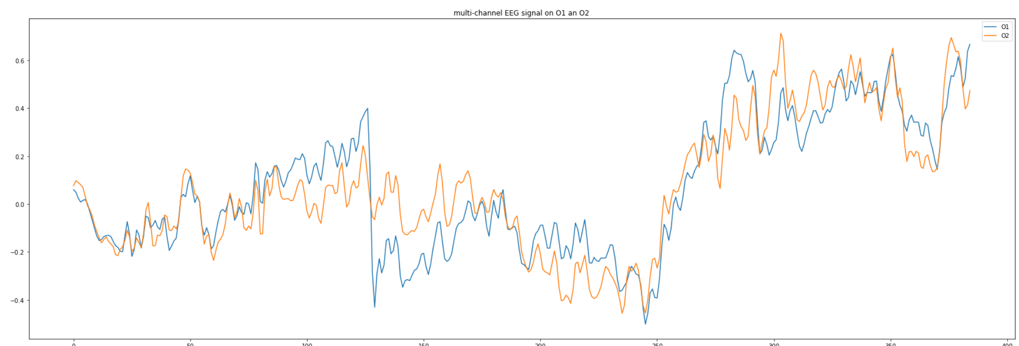

Figure 3 EEG signals in O1 and O2 channel

When we input a single O1 channel EEG signal dataset into the original CNN-LSTM, we got the training accuracy turned out to be 65%, compared to 72% when inputting the driver dataset. The training accuracy for single O2 channel EEG signal dataset is 64.63%. These results show the limitation of the original single channel neural network, and the necessity that we need to re-design the network. With this idea, we combine O1 and O2 electrodes's EEG signals to generate the 2-channel training and testing dataset, and enlarge the CNN to 2 channels. The final accuracy will be the mean of accuracy for each group. The accuracy score can directly reveal the correct rate of our prediction, and it matches the metric used in previous studied baseline model [14]. From figure 4, the final mean accuracy is 69.4%, improved by 4% than the single-channel solution.

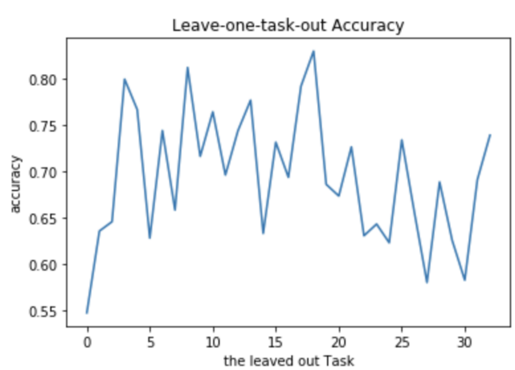

Figure 4 Leave-one-task-out Accuracy for Multi-channel CNN-LSTM

6) Compute/Other Resources Used

In this research project, we used a notebook instance with Amazon SageMaker as our main computing resources which has 10GB EBS volumes, 1 GPUs, 4 vCPUs, 61GiB RAM and high Network Bandwidth. We also use Google Colab for pre-training and program testing.

7) Conclusions

In this research, we applied a state-of-art CNN-LSTM model on EEG signal dataset for measuring the status of student attention. The model prediction accuracy is around 69.36 ± 0.1% on O1 and O2 multi-channels dataset, improved by 4% than the single-channel model, which proves our hypothesis that it is possible to use multi-channel signals for a better performance on classification tasks. We also test different activation functions, like Sigmoid and leakyReLu, and different optimizers like SGD, Softmax as activation function and Adam as optimizer still achieve the highest accuracy. According to recent research [6,17,18], there are more advanced deep learning models that can be used for analyzing and generating time series, like time GANs, N-BEATS, DeepAR and so on. Therefore, we will apply some advanced deep learning models as improvements in the future.

This model is mainly for enhancing study or work efficiency, which would impact education and the medical field. Firstly, educators can better design the homework and workload for students. With simple consumer-grade EEG devices, educators can get the brain signals of a student, send them into our model, and evaluate the concentration state to judge objectively whether the coursework is suitable. Secondly, for early-stage disease diagnoses, such as attention deficit hyperactivity disorder(ADHD)[5], our model can significantly aid the diagnosis process, and generate more valuable data insights for doctors. However, we can not ignore that when this model is widely used in the education and medical fields, a large amount of EEG data will be collected, which may raise a series of data bias and privacy issues. Data and code bias issues, on the one hand, the data we collect, especially the target value, cannot guarantee that there is no bias because whether one is concentrated or not is subjective, and it is difficult for others to judge, which in turn will lead to bias in our training results. Additionally, privacy issues should also be taken into consideration. [19] figured out that a large amount of data can be misused without authorization, and there may be some unreliable data collectors leaking information, which will create a lot of privacy issues.

(Exempted from page limit) Other Prior Work / References (apart from Sec 3) that are cited in the text:

[1]. Posner MI, Rothbart MK. Attention, self-regulation and consciousness. Philos Trans R Soc Lond B Biol Sci. 1998 Nov 29;353(1377):1915-27. doi: 10.1098/rstb.1998.0344. PMID: 9854264; PMCID: PMC1692414.
[2]. Fahimi F, Guan C, Ang K K, Goh W B and Lee T S 2017 Personalized features for attention detection in children with attention deficit hyperactivity disorder IEEE Engineering in Medicine and Biology Society (Jeju Island, South Korea) pp 414–7
[3]. C. -S. Wei, Y. -T. Wang, C. -T. Lin and T. -P. Jung, "Toward Drowsiness Detection Using Non-hair-Bearing EEG-Based Brain-Computer Interfaces," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 2, pp. 400-406, Feb. 2018, doi: 10.1109/TNSRE.2018.2790359.
[4]. M. Samavati, A. Motie Nasrabadi and M. R. Mohammadi, "Automatic minimization of eye blink artifacts using fractal dimension of independent components of multichannel EEG," 20th Iranian Conference on Electrical Engineering (ICEE2012), 2012, pp. 1576-1578, doi: 10.1109/IranianCEE.2012.6292611.
[5]. Motamedi-Fakhr S, Moshrefi-Torbati M, Hill M, Hill C M and White P R 2014 Signal processing techniques applied to human sleep EEG signals—a review Biomed. Signal Process. Control 10 21–33
[6]. Craik A, He Y, Contreras-Vidal JL. Deep learning for electroencephalogram (EEG) classification tasks: a review. J Neural Eng. 2019 Jun;16(3):031001. doi: 10.1088/1741-2552/ab0ab5. Epub 2019 Feb 26. PMID: 30808014.
[7]. Shenoy P, Krauledat M, Blankertz B, Rao R P and Muller K R 2006 Towards adaptive classification for BCI J. Neural Eng. 3 R13–23
[8]. Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, & Brent J Lance (2018). EEGNet: a compact convolutional neural network for EEG-based brain computer interfaces. Journal of Neural Engineering, 15(5), 056013.
[9]. Subasi A, Gursoy M I and Ismail Gursoy M 2010 EEG signal classification using PCA, ICA, LDA and support vector machines Expert Syst. Appl. 37 8659−66.
[10]. Messaoud, Rémy Ben and Mario Chavez. "Random Forest classifier for EEG-based seizure prediction." ArXiv abs/2106.04510 (2021): n. Pag.
[11]. A. Allahverdy, A. K. Moghadam, M. R. Mohammadi, and A. M. Nasrabadi, ''Detecting ADHD children using the attention continuity as nonlinear features of EEG,'' Frontiers Biomed. Technol., vol. 3, nos. 1–2, pp. 28–33, 2017.
[12]. N.R. Pal, et al., EEG-based subject-and session-independent drowsiness detection: an unsupervised approach, EURASIP J. Adv. Signal Process. 2008 (1) (2008), 519480.
[13]. F. Fahimi, et al., Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI, J. Neural Eng. 16 (2) 2019
[14]. Jian Cui, Zirui Lan, Yisi Liu, Ruilin Li, Fan Li, Olga Sourina, & Wolfgang Müller-Wittig (2021).
A compact and interpretable convolutional neural network for cross-subject driver drowsiness detection from single-channel EEG.
[15]. C.I. Acı, M. Kaya, Y. Mishchenko, "Distinguishing mental attention states of humans via an

EEG-based passive BCI using Machine Learning Methods", Expert Systems with Applications, vol. 134, pp. 153-166, 2019.
[16]. Schirrmeister RT, Springenberg JT, Fiederer LDJ, Glasstetter M, Eggensperger K, Tangermann M, Hutter F, Burgard W, Ball T. Deep learning with convolutional neural networks for EEG decoding and visualization. Hum Brain Mapp. 2017 Nov;38(11):5391-5420. doi: 10.1002/hbm.23730. Epub 2017 Aug 7. PMID: 28782865; PMCID: PMC5655781.
[17]. He Y, Eguren D, Azorín J M, Grossman R G, Luu T P and Contreras-Vidal J L 2018 Brain–machine interfaces for controlling lower-limb powered robotic systems J. Neural Eng. 15 021004
[18]. Treder M S, Bahramisharif A, Schmidt N M, van Gerven M A and Blankertz B 2011 Brain−computer interfacing using modulations of alpha activity induced by covert shifts of attention J. NeuroEng. Rehabil. 8 24
[19]. K.D. Martin, A. Borah and R.W. Palmatier Marketing Science Institute Working Paper Series (2016) (Report No. 16-104)