EEG alcoholism classification based on CNN with bidirectional layer *

Yukang Liu

Research School of Computer Science, Australian National University, Canberra Australia
u6920262@anu.edu.au

Abstract. Electroencephalography(EEG) is one of the most important forms that enables the visualization of brain signals modalities for neuron scientists to detect the human status. However, EEG-based status detection still remains a challenging task due to the poor interpretability of the signal and distinctiveness between subjects. However, this kind of work has made a breakthrough as there are lot recent studies involve Neural Network technique and reach the state of the art performance. In this paper, we propose a reliable framework to finish EEG-based alcoholism detection. Specifically speaking, our proposed model is based on convolution neural network(CNN) technique with bidirectional neural network(BDNN) layer. And we improve our base model through various techniques including data augmentation, shared weights, polar projection, etc. All the experiments results are tested through a public dataset UCI and the results show that our model has a comparable performance with within-subject accuracy 0.863 and cross subject accuracy 0.575, which shows potential to be further optimized in the future.

Keywords: EEG classification · Bidirectional Neural Network · Deep Learning · Alcoholism detection

1 Introduction

A brain-computer interface (BCI) system based on EEG test has played an important role in the biomedical field, and achieved great success in various kinds of researches. Brain serves as the neuron center to control the entire human-body, which is closely related to behavior, status of human-beings. That could help human give quick response facing external stimulus. Located in the different positions of the brain scalp, the EEG sensors could detect the electrical activity of the whole brain in the different signal frequencies from 0.1Hz to more than 100Hz.[1] Hence, EEG signal could compromise the data signals from the brain, which could contain rich information that could be applied in many areas, such as Alzheimer's disease diagnose detection,[2][3], emotion recognition[4], Age prediction [5].

Although success of many researches have made, EEG signal data still need a better model to extract the effective features. The difficulties are as follows: As the neural system center, the structure of brain tackles multi-tasks at the same time, which makes EEG data so complicated and filled with non-target related data. In other words, The EEG signal contains a lot of noise and it's a challenge to preserve the useful information as much as possible when cleaning the data. Besides, EEG signal tends to be individual. Different subjects vary due to their physical characteristics such as blood pressure, congenital disease, gender and age, etc. This could result in huge difference between subjects even though they are labeled in the same class in the current task. Also, normal EEG signals contain the information of sensor location and time sequence, which could be regarded as spatial and temporal information. How to integrate these features into the model is one of the critical factors to decide the performance of the model.

In recent years, deep learning starts showing its power to extract the effective features from the noisy and heavy EEG signal to further improve the performance of the model. Some study focused on using the Azimuthal Equidistant Projection(AEP) as known as Polar projection to do the transaction from the raw signal data to image to reduce the potential noise of data.[6], Another study is to treat the sensors position in the brain sphere as graph based features to integrate spatial information with the signal channels [7] An adversarial learning technique is used to remove the noise of the signal data.[8] These deep-learning based paper provide alternative ways to extract the critical features from the massively structured raw EEG data.

Inspired by those papers above, we propose a reliable model for alcoholism classification. Based on the bidirectional and convolution framework and some data augmentation techniques,

* Supported by organization Australian National University.

our model acquires good performance compared with other state-of-the-art techniques, both within subject and cross subject test. Our report contribution is as follows: First, we provide analysis of EEG dataset and comparison with original one. Second, we proposed a new reliable model based on CNN and BDNN techniques and prove that the performance of the combination outperforms any of the baselines.

This report is organized as follows. In section 2, we give a brief description our dataset and the structure of our model and some data augmentation trials we have implemented to improve our model. In section 3, we would reveal some details of our experimental analysis. In the section4, we talk then about the methods to evaluate our dataset as well as the comparison with other baseline and state-of-the-art model.

2 Method

2.1 Dataset

We evaluate our model using UCI alcoholism dataset, the EEG dataset from Neurodynamics Laboratory at the State University of New York.[13][14] Our dataset contains 11057 samples with 122 subjects, among which 77 are diagnosed with alcoholism and the others are control subjects. Each sample records the average data in one second at 256Hz with 64 electrodes.[6] The classification task is to identify whether the subject has been diagnosed with alcoholism or a control subject.The dataset also provides the different stimulus type for each trial, we would talk about how to use the stimulus type to optimize our model in the next session. And our experiments are divided into 2 sets, named within-subject and cross-subject.

Although similar, there are two different parts compared with the original dataset: One is that we reduce the time dimension by simply average the time sequence data in one trial. That means in each trial, we have 192 features instead of 64×192 matrix. Another one is that the dataset is unbalanced, not only in the label distribution, but also the distribution of subject samples.In the original dataset, each subject has 120 trials independently, we note that the number of trials each subject ranges from 30 trials to 119 in our current dataset.

Due to complexity of EEG signals and the simplicity of the current dataset. We are faced with several challenges as follows: First, since each subject has an unbalanced number of trials, the trained model tends to have bias to the subject which has more samples. In addition to the unbalanced labels, the model without any data preprocessing techniques would be easy for overfitting, especially when it comes to cross-subject testing. Second, the temporal information contains so much information that some natural language processing techniques has used it to achieve great success. [11]Brutally omitting the time dimension would result in worse performance of the model eventually. We need to provide a reliable model and some other techniques to address these difficulties.

2.2 CNN

Convolution Neural Network(CNN) is a deep learning algorithm which can compress the image-structured data and retrieve the information which makes the current image distinctive with others using convolution theory. In general, CNN consist of several different types of layers: Convolution layer is to use the convolution operation to extract the feature of current inputs. One convolution layer always contains multiple kernels with the same size, so that one matrix would be mapped into multiple smaller matrix with different kernels, the feature of the image would retrieved comprehensively. Due to all the kernel process would be linear mapping, the convolution layer also contains an activation function. Pooling layer is to further reduce the dimension and conclude the local dominant feature, instead of convolution, the pooling layer usually takes the average or maximum of the current windows as features, which could relief the computational requirements but also increase the reliability to the variation of the image, such as rotation or movement. Then when the break down features flatten dimension is simple enough for the model to take as input, the fully-connected layer is used to flatten the features. In this study, since the EEG data is transformed into the image data, a CNN-based model is needed to further improve the with-in and cross accuracy of our model.

2.3 BDNN

Bidirectional Neural Network provides a new way to improve the network converging efficiency through training the the network backward and forward simultaneously. Figure 1 compares the network topology difference between BDNN and conventional NN model. In the conventional neural network based model, the weight updating depends on the back propagation mechanism using the loss between the predicted output and its corresponding target. However, the training could be seen in the reverse direction, that is, we take the original output as input, to predict the original input. Also, due to the weights them-self do not have characteristic of direction. Using shared weight technique, the weight back propagation could also use the loss between the predicting input and the true inputs and continue updating in the reverse direction. We could use the shared weight technique to finish the BDNN implementation. The relationship between the shared weights in the reverse directions are as follows:

$$y = Ax + b \tag{1}$$

$$x = A^T(y - b) \tag{2}$$

Where x, y, A and b represent the input, output, the weights, bias(could be dropped) respectively. With simple structure of neural network, a more complicated network topology is needed to retrieve the information from the data. Hence, it is necessary to propose the BDNN instead of normal NN-based model to tackle with the classification problem.

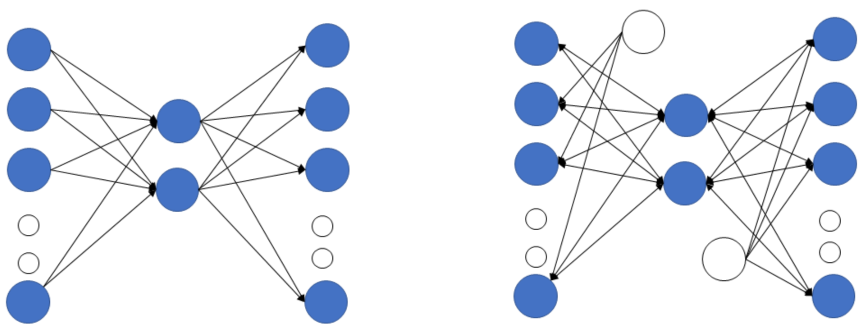

Fig. 1. comparison between NN and BDNN within one layer

2.4 Data Preprocessing

As we state in the introduction session, the dataset is unbalanced both in the labels distribution and subject trials. Here we tried several ideas to improve the quality of our raw data.

EEG-to-Image projection Inspired on the Bashivan's work[12], a transformation from EEG to RGB encoding data is used in our study: Since the our dataset has took the average of time-series each channel, we combined the the 3 channel(frequency) EEG data with location of the senors in the brain. Using Azimuthal Equidistant Projection as known as polar projection algorithm, the 3-D sensor position would be mapped into 2-D space, which helps EEG data construct its corresponding 32x32x3 image. In such, the EEG data would be visualized as RGB-encoding image, which provides chance for us to apply CNN-related model in the later study.

Input Normalization Due to the distinctiveness of subject, the features vary different between subjects to subjects. Even in the same subject, different stimulus result in various EEG signal. Figure 2 illustrate both inner-subject and cross-subject difference. In the left table, all the samples are derived from subject 111, which should belong to the same class for all the samples, but with different stimulus, the data distribution varies. In the right graph, the huge difference exists between subject even though all of the 3 subjects are diagnosed with alcoholism and are tested with the same stimulus.

This essential part of neural network is to learn the distribution of dataset. On the one hand, the features ranging differently would result the different distribution of the data every batch, which would slow down the training process. On the other hand, when the testing and training data have different distribution, the generalization of the model would be affected. Hence, we adopt input normalization in our preprocessing procedure.

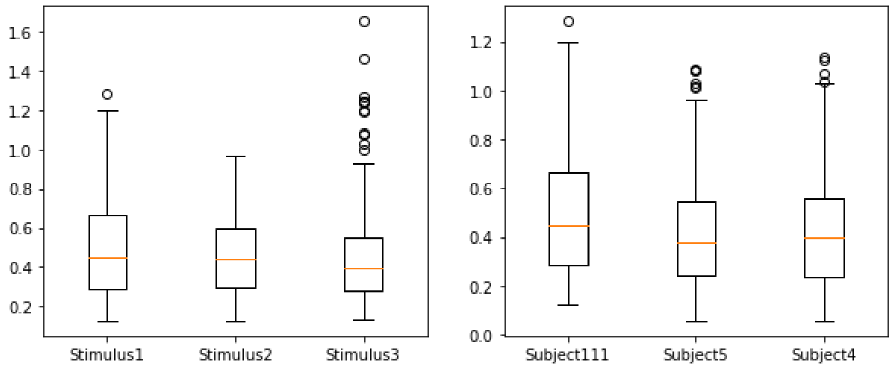

Fig. 2. Data inner and cross distribution difference

Enhanced Class The binary classification is kind of simple, but really hard to distinguish the samples that are closed to the border. The complexity of EEG signal enhanced this problem since its poor interpretability so that we cannot sure. To address this problem. we introduce enhanced classes: Since each trial the subject receives a specific stimulus. We integrate the 5 stimulus into the status of alcoholism and consist of 10 classes labeled from 0 to 9, among which 0-4 are control label and the other is the control subjects.Then we turn the model into tackling with 10 labels classification problem instead and evaluate the model accuracy by summing up the labels which should fall into 0-4 or 5-9 originally.

2.5 model optimizing

Fine-tuning is an essential work to train an model. We do some model optimizing work as follows:

Optimizer choosing We choose Adam[10] as our model optimizer since Adam can do well with a data-set full of noise. Adam combines the property of algorithm Adamgrad[9] and RM-Sprop[10], which compromises first moment estimation and second moment estimation to decide next update weight. Compared with normal SGD, the Adam optimizer could update learning rate automatically and is not easy to be affected by the gradient changing. Hence, the Adam optimizer is more robust when noisy dataset comes in.

To prove the efficiency of Adam optimizer, we compare the optimizer with SGD and SGD with momentum. (shown in figure 3)

hidden units and multi hidden layers One of the effective way to increase the structure of complexity is to changing the hidden units. Here we list part of our tests as our results. With the same other setting, the training loss comparison are as follows(see figure 4)

2.6 structure of our models

Combining what we have stated above, our model pip-line is defined as follows: First we pre-process our data using input normalization and enhanced class mechanism, then using polar

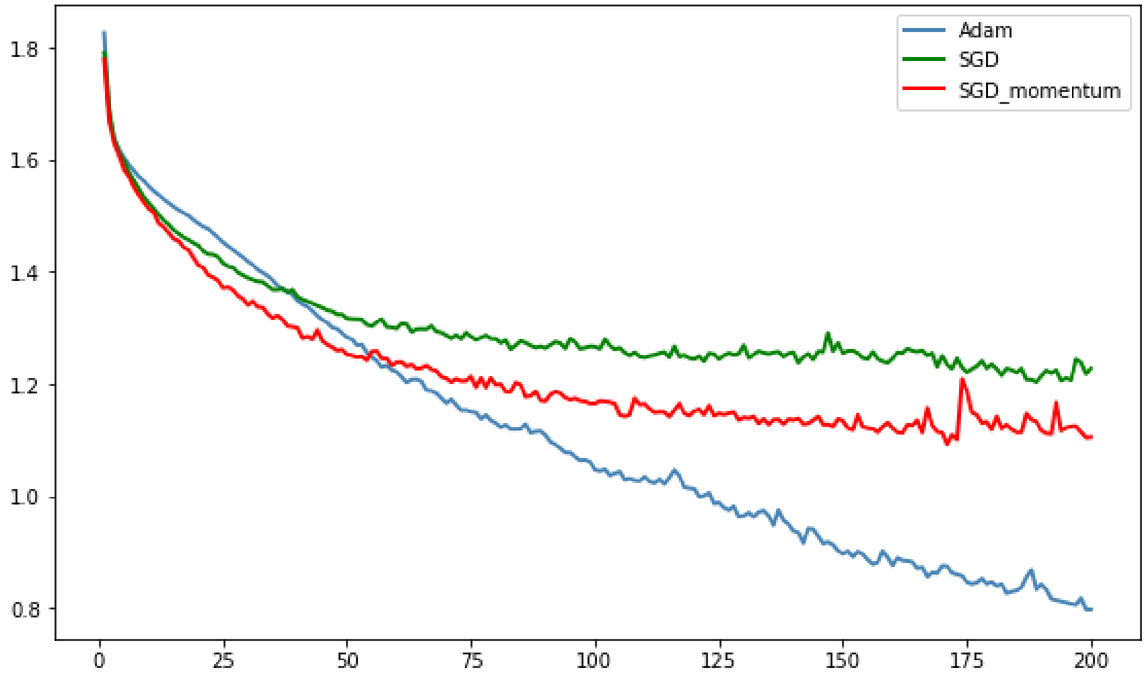

Fig. 3. Data inner and cross variety

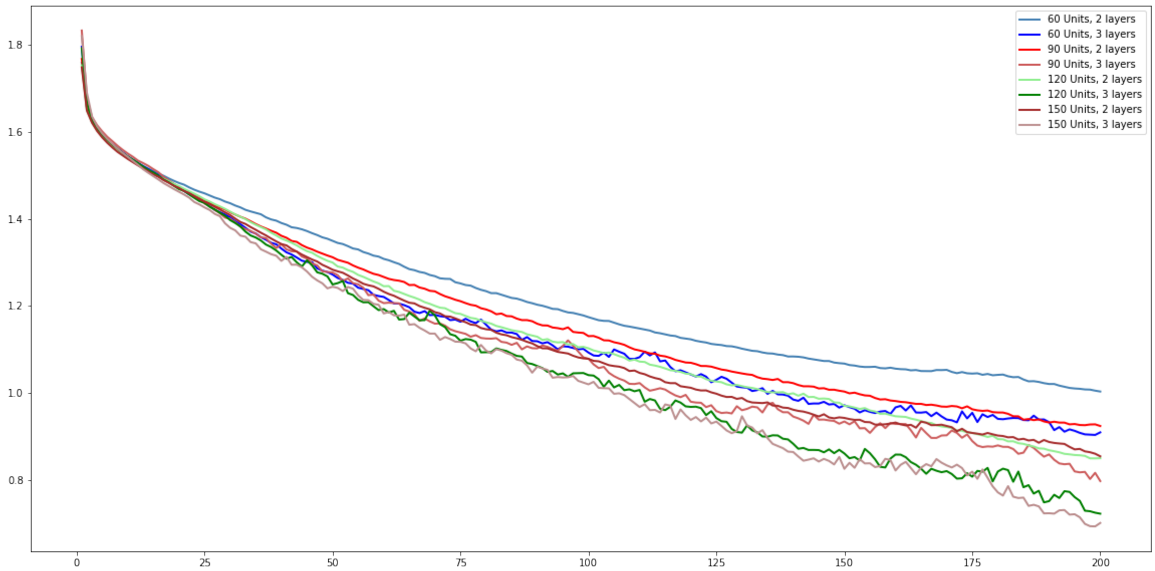

Fig. 4. Data inner and cross variety

projection and CloughTocher scheme to transform the EEG data into image and split the data into training set, test set and validation set. Then the train data would help train our proposed CNN model with one BDNN layer. Regarding implementation, our model is constructed as figure 5 : it consists of three pairs of convolution-max-pooling layers, converting the data with the size of 3x32x32 into 128x4x4. Then the model would flatten the compressed data into a vector with the size of 2048, then it would go through a BDNN layer instead of fully-connected layer. BDNN contains two conventional neural nets, one called forward net and the other called feedback neural net. Both neural nets consists of 1 input layer, 2 hidden layers and 1 output layer but the input and output units are reversed.

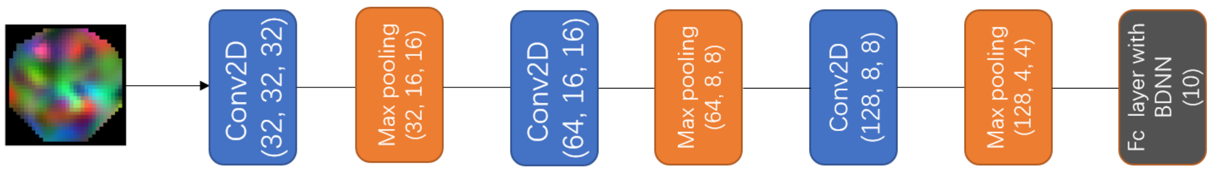

Fig. 5. Model pipelines

3 Results and Discussion

3.1 Evaluation Method

Our implementation is written in python and pytorch. All experiments are tested in the windows environment and an i5-7300HQ, Nvidia GTX 1050, 12g RAM.For evaluation method, we split our dataset into train set, test set, validation set with the size proportion 7:2:1. The train set is used to train the model, test set is used to fine-tuning. The validation set evaluates the final model performance.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$precision = \frac{TP}{TP + FP} \tag{4}$$

$$recall = \frac{TP}{TP + FN} \tag{5}$$

where TP, TN, FP, FN represent true possible, true negative, false positive, false negative respectively. As we have mentioned in the introduction section. We divide our experiments into 2 sets: within-subjects and cross-subjects. Within-subjects refers to the model is trained by the data for one specific subject, while cross-subject means the model is trained by part of subject data and test the performance through those subjects who do not involve in the training set data. Due to the difference between subjects, the cross subject appeal to be lower performance but more challenging. Since our dataset samples are not distributed evenly. We decide to use the 5 subjects who has the most samples do the inner-subject experiments, and average the validation accuracy as the model final performance.

3.2 Comparison with base line

Due to the difference between the dataset, we decide to set up our model baseline initially to prove our model performance. There are three base line models. One is the simple neural network with one hidden layer. inspired by the idea of Yao's paper[6], we also set up a CNN model using polar project and FFT mechanism to process the data and then use a basic CNN

Table 1. comparison between proposed model and baseline

	CNN	BDNN	NN	CNN + BDNN
Within-subject train accuracy	0.976	0.91	0.821	**0.94**
Within-subject validation accuracy	0.814	0.850	0.843	**0.863**
Cross-subject train accuracy	0.579	0.604	0.416	**0.621**
Cross-subject validation accuracy	0.421	0.572	0.319	**0.575**

Table 2. comparison between proposed model and state-of-the art

method	within accuracy	cross accuracy
CNN + BDNN	0.863	0.575
Normal Channel-wise Autoencoders	0.864	0.731
Normal Image-wise Autoencoders	0.917	0.756
EEGNet	0.878	0.672
SyncNet	0.923	0.723
PSD	0.816	0.605

model containing 6 convolution layers, 2 pooling layers and 2 fully-connected layers. Besides, based on previous attempt, we set a normal BDNN as a baseline as well.

In the first table, we can see that the combination of CNN and BDNN outperforms any of the single type of the neural network. In the within-subject test, all of them reaches above 0.8 validation accuracy. However, one thing to be noted is that the huge difference between train accuracy and validation accuracy in the CNN model tends to be overfitting, which mainly because the structure of CNN does not take many procedure to reduce the noise of the dataset. In contrast, a simple 2 layer neural network seems to simple to distinguish the difference between subjects, resulting in the low accuracy when it comes to cross-subject tests. Moreover, the combination of the two neural networks shows better performance compared with the other baselines, which may because compared a single neural network, bidirectional neural network has a positive effect to stop the model prune to overfitting.

However, compared with the state-of-the-art models, the result shows that our model needs to be further modified to reach the same-level accuracy. On the one hand, our model reaches 86 accuracy in the within-subject test. The reason behind this may be the size of trials each subjected is restricted within 120, which is really simple for the model to train. On the other hand, the model performs worse than any other state of the art models in the cross-subject testing. We believe this is caused by the lack of the dataset, including the missing samples and time sequence information, besides there are a lot of hyper-tuning work left to do to optimize our model, even though our proposed model has a less running time due to the simple structure: our proposed model could finish the whole classification work within 100 seconds with 100 epoch, compared with 132.99s and 150.68s with yao's idea [6]

4 Conclusion and Future Work

In this report, we proposed a framework based on bidirectional neural network and convolution neural network to address the alcoholism detection from EEG signal. The results shows our proposed neural network has shown better performance in comparison with simple neural network, single CNN and single BDNN and reach a state-of-the-art accuracy performance in the within-subject test. However, the model does not bring the good results when it comes to the cross-subject test due to the fact that the the structure of our model is simple and only part of the dataset are being used. In the realistic application, we are often required to predict the alcoholism status using the EEG signal from an unseen subject. Hence, it's more important for us to explore more optimization methods to improve the cross-subject performance further.

Besides, there are many data not used in the UCI dataset, we might try to use the orignal entire dataset which contains time series data and use deep learning model such as LSTM to integrate spatial and temporal information as our features instead of brute taking means of the time record data. Then, the future work could also include more data augmentation method to balance class weight and remove the noise of the EEG signal without losing most of the features. In the end, since we would have many hyperparatmeters to fine-tune in the later deep learning technique implementation, a good tuning strategy such as one cycle policy need to be done to improve the model efficiency.[15]

References

1. J. Satheesh Kumar and P. Bhuvaneswari.: Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study. Procedia Engineering. **38** ,2525-2536 (2012)
2. Dauwels J, Vialatte F, Cichocki A.:Diagnosis of Alzheimer's disease from EEG signals: where are we standing? Curr Alzheimer Res. 2010 Sep;7(6):487-505. doi: 10.2174/156720510792231720. PMID: 20455865.
3. Gao Yunyuan, Gao Bo, Chen Qiang, Liu Jia, Zhang Yingchun.:Deep Convolutional Neural Network-Based Epileptic Electroencephalogram (EEG) Signal Classification. Frontiers in Neurology, **11**, 375(2020)
4. Zheng W L, Lu B L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks[J]. IEEE Transactions on Autonomous Mental Development, 2015, 7(3): 162-175.
5. Al Zoubi O, Ki Wong C, Kuplicki R T, et al. Predicting age from brain EEG signals—A machine learning approach[J]. Frontiers in aging neuroscience, 2018, 10: 184.
6. Yao Y., Plested J., Gedeon T.: Deep Feature Learning and Visualization for EEG Recording Using Autoencoders. In: Cheng L., Leung A., Ozawa S. (eds) Neural Information Processing. ICONIP 2018. Lecture Notes in Computer Science, vol 11307. Springer, Cham. https://doi.org/10.1007/978-3-030-04239-450 (2018)
7. Zhong P, Wang D, Miao C.: EEG-based emotion recognition using regularized graph neural networks[J]. IEEE Transactions on Affective Computing, (2020)
8. Y. Yao, J. Plested, T. Gedeon, Y. Liu and Z.:Wang, Improved Techniques for Building EEG Feature Filters. International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-6, doi: 10.1109/IJCNN.2019.8852302.(2019)
9. Duchi, J., Hazan, E., Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. Journal of machine learning research, 12(7).(2011)
10. Ruder, S.: An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.(2016)
11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I.:Attention is all you need. arXiv preprint arXiv:1706.03762.(2017)
12. Bashivan, P., Rish, I., Yeasin, M., Codella, N.: Learning representations from EEG with deep recurrent-convolutional neural networks. arXiv preprint arXiv:1511.06448.(2015).
13. Peter Sykacek and Stephen J. Roberts. Adaptive Classification by Variational Kalman Filtering. NIPS. 2002.
14. Stephen D. Bay and Dennis F. Kibler and Michael J. Pazzani and Padhraic Smyth. The UCI KDD Archive of Large Data Sets for Data Mining Research and Experimentation. SIGKDD Explorations, 2. 2000.
15. Smith, L. N.: A disciplined approach to neural network hyper-parameters: Part 1–learning rate, batch size, momentum, and weight decay. arXiv preprint arXiv:1803.09820. (2018)