

EEG-Signals Based Cognitive Workload Detection of Vehicle Driver using Deep Learning

Mohammad A. Almogbel*, Anh H. Dang**, Wataru Kameyama***

*Department. of Computer Science. and Communications Engineering
Graduate School of Fundamental Science and Engineering, Waseda University, Tokyo, Japan.

**GITS, Waseda University, Tokyo, Japan

***Faculty of Science and Engineering, Waseda University, Tokyo, Japan

almogbela@ruri.waseda.jp, danghoanganh@akane.waseda.jp, wataru@waseda.jp

Abstract— Vehicle driver's ability to maintain optimal performance and attention is essential to ensure the safety of the traffic. Electroencephalography (EEG) signals have been proven to be effective in evaluating human's cognitive state under specific tasks. In this paper, we propose the use of deep learning on EEG signals to detect the driver's cognitive workload under high and low workload tasks. Data used in this research are collected throughout multiple driving sessions conducted on a high fidelity driving simulator. Preliminary experimental results conducted on only 4 channels of EEG show that the proposed system is capable of accurately detecting the cognitive workload of the driver with an enormous potential for improvement.

Keywords— Deep Learning, EEG, Neural Networks, Cognitive Workload, Driving, Stress

I. INTRODUCTION

Maintaining safe levels of cognitive workload is extremely crucial to ensure optimal performance and attention whilst driving automobiles. Different methods have been widely investigated to monitor driver's cognitive workload including heart rate monitoring, galvanic skin response, facial expression and so on. One of the most popular measures for assessing the mental state of humans is electroencephalography (EEG) signals. EEG has reportedly shown success in evaluating the mental state such as drowsiness [1], mind wandering [2] and alertness [3].

Despite being an active research area, understanding and applying EEG are still limited due to the lack of true understanding of the brain's activities which prevents use of good features from being engineered. Besides, EEG signals are strongly affected usually by noise and interference. Accurate reading requires delicate equipment and sealing.

Conventional analysing approaches make extensive use of Fourier transform in order to decompose EEG signal into multiple frequencies. Only well-known frequency is then used for feature design process. For example, the alpha band (8 to 15[Hz]) is correlated with relaxing, and the beta band (16 to 31[Hz]) associates with mental stress. Pre-processing methods such as Butterworth bandpass filter and stationary wavelet transform filters [4] are used to remove high and low frequency noises [3]. Several efforts in applying deep learning on EEG have been carried out [1, 8, 9, 10]. However, none of the

previous researches come up with a system that can handle raw data directly. Thus, valuable information might be discarded during the pre-processing.

The purpose of this paper is to introduce an end-to-end deep neural network model that can directly infer cognitive workload from raw EEG, where signal pre-processing as well as conventional feature design are not required.

To evaluate the proposed model, EEG signal recordings have been carried out on a subject driving a vehicle in a relatively realistic simulated environment under two different types of cognitive workloads, high workload and low workload.

This paper is organized as follows: Section II explains the model used to classify the data, Section III introduces how the data are collected and prepared, Section IV evaluates how the model is trained with the results of the proposed model, and the conclusion is given in Section V.

II. MODEL

One of the major advantages of convolutional neural networks (CNN) is the ability to extract, learn and generalize features without pre-processing. In this study, we propose the use of deep CNN to predict the cognitive workload of the driver by using raw EEG signals with no additional pre-processing.

A. EEG

Monitoring electrical brain activity in a non-invasive means is referred to as EEG in clinical context. Brain activities produce electrical charges caused by the neurons inside the brain, and voltage fluctuations are measured using metal pieces called electrodes and a reference electrode attached to the head and scalp [5]. These voltages pass through an amplifier for analysis. Figure 1 shows the international 10-20 system that is used to describe the location of the electrodes on the head. Thanks to recent advancements, compact lightweight devices have been introduced into the market to allow a more consumer friendly means to monitor EEG signals.

In this research, EEG recording has been done using Muse [6] headband. Figure 1 also explains the electrodes used in this device where 4-channel configuration has been utilized; two dry electrodes on the forehead AF7 (Left) and AF8 (Right), two behind the ears TP9 (Left) and TP10 (Right), and a reference electrode (Fpz) in the middle of the forehead above the nasion.

This device is capable of recording EEG signals at a sampling rate of 256[Hz]. Moreover, it is very portable, lightweight, easy to use, and it utilizes low-energy Bluetooth technology to transmit data making it feasible for real-life usages.

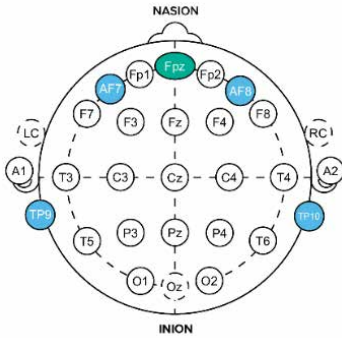


Figure 1. Locations of Scalp Electrodes in International 10-20 System [6]

B. Network Architecture

CNN is inspired by biological visual system [7] where different levels of visual hierarchy are handled by different parts of the brain. Analogically, multi-layers of CNN generalize visual features in multiple levels. Since each layer of CNN acts as a filter, it is expected that CNN can be also used to extract high level features even in the case of EEG signal.

Different from images, EEG comprises multiple time series from each channel. For that reason, it is essential to design CNN filters to stride along time direction only. Thus, it must fully cover all other dimensions of input data.

The architecture of the proposed network is visualized in Figure 1. As shown in the figure, in the case of 150[sec] of EEG data as the input window size, it consists of 37500 data samples for each channel, thus the raw EEG data of 37500×4 , is used to the input. The optimum input window size is discussed in Section IV. Then, seven convolution layers transform raw EEG data down to 3×8 features. These features are then flattened and processed by 3 layers of fully connected (FC) classifier block. All CNN filter strides along time dimension.

In a proposal such as [9], separated CNN networks are used on each band of EEG channel independently. However, our CNN filters are applied to all EEG channels in the same time. This method ensures that all possible combination features are captured. Small filter size and stride together with high number of CNN layers also ensure that features from multiple hierarchical levels are extracted.

The FC block outputs a two-dimension vector. Softmax function is used to interpret this vector into the probability of each cognitive workload, one for high cognitive workload and one for low cognitive workload. Except the final layer of FC block, all layers in this deep network utilize rectified linear unit (RELU) [8] activation function.

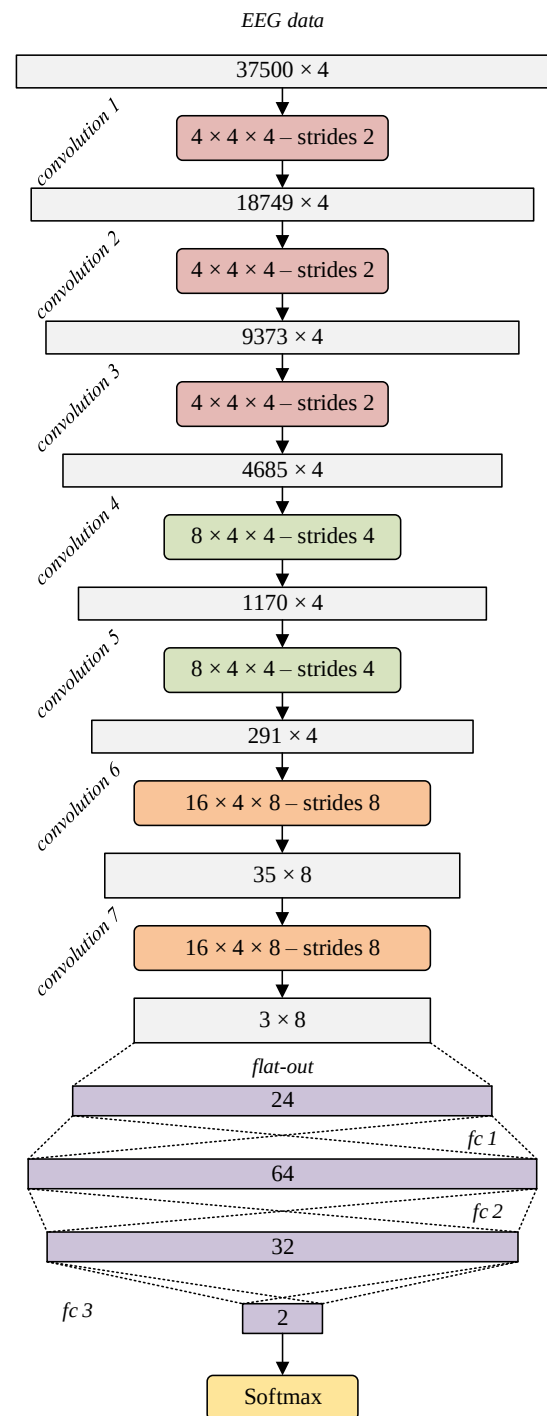


Figure 2. Proposed Deep Neural Network Model in the case of 150[sec] of EEG Input (37500 Samples in Each Channel)

III. DATA COLLECTION

Experiment recordings have been conducted on a popular driving simulation video game (GTAV free driving mode) that is highly capable of representing real life driving situations. Other used equipment are Logitech G27 driving wheel and an ultrawide (21:9) curved monitor as shown in Figure 3.



Figure 3 Representation of the Used Simulation Environment with an Example of High Workload (Left) and Low Workload (Right)

A. Driving Sessions

For the purpose of evaluating the cognitive workload of the subject, we choose two different types of driving sessions. The first type imposes high cognitive workload, i.e. dense traffic situation and complex metropolitan scene, and the opposite for the second type, i.e. empty roads and monotonous highway. The subject is advised to drive as if they would in actual situation. We implicitly assume that driving in the first type is more mentally demanding than the second type.

Each of the two types of session was recorded for an interval of 15 to 30 minutes on a span of one month. Data collections were randomly done in daytime and night-time with random intervals between them, and each session had a different random route to remove any bias. The subject was required to disengage from the test-bed upon completing each session.

In this experiment, one subject of male with age 29 was joined, who drives frequently with 11 years of driving experience. The number of sessions for the experiments is 24 where 12 is for high workload and 12 is for low workload.

B. EEG Sampling Rate

Data were collected at 256[Hz] sampling rate by Muse as mentioned in Section II A. Hence, the data in one second consist of 256 samples, and that of 15 to 30 minutes session consist of 230,000 to 460,000 samples for each of the 4 channels.

IV. EVALUATION

A. Data Preparation

The data of all the sessions are used for training the model except for that of the last session which is used for the evaluation. By doing so, real-life situation, where the model requires to estimate the workload from future data by learning previous data, can be simulated. Thus, over 6 million samples for each of the 4 channels are used for the training and 460 thousand for the evaluation.

In order to train our proposed model, raw EEG data are normalized by z-score for each of the 4 channels. Slicing window approach is utilized to prepare the data for training. Table 1 explains the different slicing window size used for training the model as well as the shape of the input data, where each slice is denoted by the time segments of the EEG data. The slicing windows are overlapped with 1/256[sec] step.

B. Experiment

1) Training the model: The model is trained for a total of 10 epochs using RMSProp optimizer with a learning rate of 0.002. Each epoch represents the training of the model throughout all the data. Input data are divided into batches of 64. During this training process, 50% of drop-out is applied to all layers (except the last layer) within FC block. Figure 4 shows the learning curve of the best resulting window size of 150[sec] alongside with the evaluation curve.

2) Evaluating the model: For every iteration of the training process, random batches of the test data and their corresponding labels are used for cross-validating to evaluate the model's performance. Figure 4 also shows the evaluation where a slight overfitting can be observed towards the end of the training. This can be potentially resolved by collecting and adding more data into the training process.

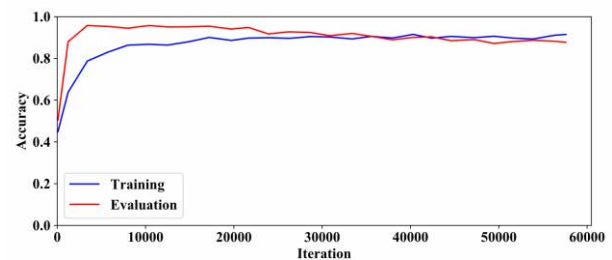


Figure 4. Learning Curve for Training and Evaluation of the Model with the Best Window Size (150 [sec])

C. Results and Discussion

The proposed model in this study is composed of 7 CNN layers that can figure the necessary transformations to extract features, and 3 fully connected layers that can accurately classify the cognitive workload of the driver. This end-to-end network utilizes only raw EEG signals from 4 channels as it's input. Table 1 summarises the achieved evaluation accuracies when running the model on different window sizes of the raw EEG data. As shown in the table, 150[sec] window size achieves the best accuracy in classifying the cognitive workload of the driver between high and low workloads. As an additional metric to evaluate our model, classification accuracy of high cognitive workload is 95.76%, and 92.57% for low cognitive workload sessions.

The proposed network is originally designed and optimized for 150[sec] window size. Therefore, there could be any other optimum configurations in the network structure for other window sizes. In this experiment, the 7th convolutional layer was removed for the 30 and 60 [sec] window size.

TABLE 1 EEG DATA WINDOW SIZES USED TO EVALUATE THE MODEL AND THE CORRESPONDING EVALUATION ACCURACIES.

EEG Data Window Size	Input Shape	Evaluation Accuracy
30 Seconds	7680×4	87.25%
60 Seconds	15360×4	92.02%
90 Seconds	23040×4	88.38%
120 Seconds	30720×4	82.24%
150 Seconds	38400×4	95.31%
180 Seconds	46080×4	84.69%

Conventional approaches to classify EEG signals consist of decomposing the signals to extract features to be used for classification. Attempts of EEG analysis to utilize deep learning with minimal pre-processing are shown in Table 2. [9] evaluates drivers' cognitive performance in a simulated environment using filtered frequency EEG. In [1], FFT is used before classifying the signals to detect driver's drowsiness, and [10] also proposes a model that can predict right and left hand movements with frequency filtered EEG signals. [11] apply a spatial filter before classifying pathological from normal EEG recordings.

This study does not impose in any way a direct comparison with the distinguished previous works because the used data, experimental conditions, classification targets are different in each, but rather explore and introduce the potential of using deep CNN architecture in classifying raw EEG signals without any pre-processing.

TABLE 2 RECENT RESEARCHES ON EEG ANALYSIS WITH DEEP LEARNING

Ref.	Pre-processing	Feature-Extraction	Classifier	Accuracy
[9]	Frequency filtering	CNN	1-layer ANN	86.06%
[1]	FFT	N/A	1-layer ANN	86.50%
[11]	Spatial Filter	CNN	1-layer ANN	84.80%
[10]	Frequency filtering	CNN	1-layer ANN	86.41%

V. CONCLUSION

In this research, we propose an end-to-end deep neural network that can accommodate raw EEG signals from 4 channels collected within one month from numerous driving sessions as its input. The results show that the proposed model is capable of successfully generalizing the EEG signals and making highly accurate classification of the driver's cognitive workload.

Future works include testing the proposed model with public available data sets. Further investigation will be carried out by collecting more data from more subjects with different driving experiences and classifying more than two types of cognitive workloads.

REFERENCES

- [1] I. Belakhdar, W. Kaaniche, R. Djmel, and B. Ouni, "Detecting driver drowsiness based on single electroencephalography channel," *13th Int. Multi-Conference Syst. Signals Devices, SSD 2016*, pp. 16–21, 2016.
- [2] C. L. Baldwin, D. M. Roberts, D. Barragan, J. D. Lee, N. Lerner, and J. S. Higgins, "Detecting and Quantifying Mind Wandering during Simulated Driving," *Front. Hum. Neurosci.*, vol. 11, no. August, pp. 1–15, 2017.
- [3] L. Bi, R. Zhang, and Z. Chen, "Study on Real-time Detection of Alertness Based on EEG," *2007 IEEE/ICME Int. Conf. Complex Med. Eng.*, pp. 1490–1493, 2007.
- [4] S. S. Daud and R. Sudirman, "Butterworth Bandpass and Stationary Wavelet Transform Filter Comparison for Electroencephalography Signal," *Proc. - Int. Conf. Intell. Syst. Model. Simulation, ISMS*, vol. 2015–October, pp. 123–126, 2015.
- [5] S. Donald L. and F. H. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2011.
- [6] Interaxon, "Muse: the brain sensing headband," *Tech. Specif. Valid. Res. use*, pp. 4–9, 2017.

- [7] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, 1980.
- [8] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," *Proc. 27th Int. Conf. Mach. Learn.*, no. 3, pp. 807–814, 2010.
- [9] M. Hajinoroozi, Z. Mao, T. P. Jung, C. T. Lin, and Y. Huang, "EEG-based prediction of driver's cognitive performance by deep convolutional neural network," *Signal Process. Image Commun.*, vol. 47, pp. 549–555, 2016.
- [10] Z. Tang, C. Li, and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik (Stuttgart)*, vol. 130, pp. 11–18, 2017.
- [11] R. T. Schirmer, L. Gemein, K. Eggensperger, F. Hutter, and T. Ball, "Deep learning with convolutional neural networks for decoding and visualization of EEG pathology," 2017.



Mohammad A. Almogbel (S'14) received his bachelor's degree in Information Systems from King Saud University, Riyadh, Saudi Arabia in 2009. He joined King Abdul-Aziz City for Science and Technology in Saudi Arabia as a researcher in the same year and received a scholarship to complete his graduate school in 2010. He then received master's degree in computer science from Waseda University in 2014 and he continued to pursue his Ph.D. since then. He is a member of IEEE, ITS and JSAE.



Anh H. Dang (S'09) received his bachelor degree in business administration, information & communication technology from Ritsumeikan Asia Pacific University (Beppu, Oita, Japan) in 2010. He then received the master degree in computer science from Waseda University (Shinjuku, Tokyo, Japan) in 2012. Since 2012, he is a Ph.D. candidate at Waseda University. He is a member of IEEE, ACM, and IEICE. His research interests are machine learning, artificial intelligence, and computer vision.



Wataru Kameyama (M'86) received the bachelor's, master's, and D.Eng. degrees from the School of Science and Engineering, Waseda University, in 1985, 1987, and 1990, respectively. He joined ASCII Corporation in 1992, and was transferred to France Telecom CCETT from 1994 to 1996 for his secondment. After joining Waseda University as an Associate Professor in 1999, he has been a Professor with the Department of Communications and Computer Engineering, School of Fundamental Science and Engineering, Waseda University, since 2014. He has been involved in MPEG, MHEG, DAVIC, and the TV-Anytime Forum activities. He was a Chairman of ISO/IEC JTC1/SC29/WG12, and a Secretariat and Vice Chairman of the TV-Anytime Forum. He is a member of IEICE, IPSJ, ITE, IEEE, and ACM. He received the Best Paper Award of Niwa-Takayanagi in 2006, the Best Author Award of Niwa-Takayanagi in 2009 from the Institute of Image Information and Television Engineers, and the International Cooperation Award from the ITU Association of Japan in 2012.