

A Driving Performance Forecasting System Based on Brain Dynamic State Analysis using 4D Convolutional Neural Networks

Chin-Teng Lin¹, *Fellow, IEEE*, Chun-Hsiang Chuang^{2*}, *Member, IEEE*, Yu-Chia Hung¹, Chieh-Ning Fang¹, Dongrui Wu³, *Senior Member, IEEE*, Yu-Kai Wang^{1*}, *Member, IEEE*

¹Australia Artificial Intelligence Institute, FEIT, University of Technology Sydney, Sydney, Australia

²College of Education, National Tsing Hua University, Hsinchu, Taiwan

³School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China

Abstract— Vehicle accidents are the primary cause of fatalities worldwide. Most often, experiencing fatigue on the road leads to operator errors and behavioral lapses. Thus, there is a need to predict the cognitive state of drivers, particularly their fatigue level. Electroencephalography (EEG) has been demonstrated to be effective for monitoring changes in the human brain state and behavior. Thirty-seven subjects participated in this driving experiment and performed a perform lane-keeping task in a visual-reality environment. Three domains, namely, frequency, temporal and 2D spatial information, of EEG channel location were comprehensively considered. A 4D convolutional neural network (4D CNN) algorithm was then proposed to associate all information from the EEG signals and the changes in the human state and behavioral performance. 4D CNN achieves superior forecasting performance over 2D CNN, 3D CNN and shallow networks. The results showed a 3.82% improvement in the RMSE, a 3.45% improvement in the error rate, and a 11.98% improvement in the correlation coefficient with 4D CNN compared with 3D CNN. The 4D CNN algorithm extracts the significant theta and alpha activations in the frontal and posterior cingulate cortices under distinct fatigue levels. This work contributes to enhancing our understanding of deep learning methods in the analysis of EEG signals. We even envision that deep learning might serve as a bridge between translation neuroscience and further real-world applications.

Index Terms—EEG, driving, response time, deep learning, convolutional neural network

I. INTRODUCTION

Driving safety is an important public issue. According to the World Health Organization, over 1.2 million people worldwide die in car accidents each year, while 20 to 50 million people are nonfatally injured in car accidents [1]. Lapses in attention and other behavioral errors of drivers are responsible for 90% of such accidents [2]. If this urgent issue is not improved, road traffic injuries could become the seventh leading cause of death across the world by 2030 [3]. Among these issues, this study focuses on driving while drowsy, which results in a considerable number of road traffic fatalities and

nonfatal injuries every year. There is a need to monitor drivers' fatigue levels to control and reduce road traffic injuries.

Previous studies have demonstrated that drowsy driving is highly correlated with behavioral lapses before accidents [4]–[7]. There are two major approaches for monitoring drowsy driving. One approach targets the behavior of drivers by detecting the movement of the steering wheel or the deviation of the vehicle [8]. The other approach is to monitor physiological parameters, such as the driver's heart rate, breathing, brain waves and respiratory rate [9]. Electroencephalography (EEG) has been shown to be a reliable indicator of human cognitive states. Many EEG-based driver assistance systems [8]–[15] have been proposed for observing the driving behaviors or fatigue levels of drivers. Some systems apply statistical models based on knowledge from previous research. The others apply machine learning methods, such as support vector machines (SVM) [10], Gaussian mixture models [16] and neural networks (NN) [13]. These studies suggest that EEG is a feasible tool for monitoring driving behavior by connecting the driver's cognitive state with response time.

Recently, several deep learning methods have exhibited high research value in many fields. Especially, the capacity of auto-feature learning shows great success in computer vision society [23], [24]. Compared with shallow algorithms, deep learning algorithms learn the high-level features from raw data by using multiple stacked layers. Recently, several EEG studies have adopted deep learning algorithms [42–46]. Liu et al. applied a convolutional neural network (CNN) to motor imagery tasks [25]. Hajinoroozi et al. tried to predict drivers' cognitive states (drowsy or alert) with a channel-wise convolutional neural network (CCNN) [26]. Lu et al. utilized restricted Boltzmann machines (RBM) for motor imagery tasks [27]. The researchers leverage the power of deep learning to explore the differences of brain dynamics as performing specific tasks or experiencing changes of cognitive state.

However, how to better leverage the knowledge of EEG signals for different models has become a problem. In previous studies [17]–[20], frequency bands of EEG signals were shown to reflect psychological states [17], [18]. Dissanayaka's study

showed that the power of the alpha band increased when the participants were drowsy [19]. Temporal information is also essential in the EEG analysis; connections of adjacent timeframes represent a trend in brain dynamic states [21]. Spatial information from EEG channels has received attention in recent studies. Studies have shown that the cognitive workload is strongly connected with the frontal and parietal lobes [35], [38]. In Hooi's study, EEG signals were converted into a topographical image (topo-image), which is one method used to combine the electrode positions into an EEG readout [22]. Previous studies [40], [41] showed various techniques for processing multidimension data for cognitive workload assessment. Another approach for simultaneously processing multi-information is 3D CNN [23]. In 2013, a 3D CNN algorithm was proposed to process video data [23], [33]. These studies performed three-dimensional convolution and effectively extracted motion features embedded in the spatial and temporal domains. The performance of 3D CNN is promising over many applications.

To better use the knowledge from the three domains, a CNN-based deep learning algorithm named 4D CNN was applied to predict fatigue levels during a driving task. The strategy used for the utilization of spatial information as a topo-image in the analysis helps narrow down the position of significant features in the brain-lobe view. Temporal information is also preserved for tracking the signal trends. With this advantage, we are able to monitor the activated brain area over time. In this study, response times (RTs) were recorded to evaluate the cognitive performance of drivers, and the system took 6-second baseline EEG signals as the input data to predict the RTs.

II. MATERIALS AND METHODS

A. Experimental Design

The data used in this study were recorded at the Brain Research Center (BRC), National Chiao Tung University (NCTU), Hsinchu, Taiwan. The experimental task was lane-keeping. The experiment was conducted in a 360-degree

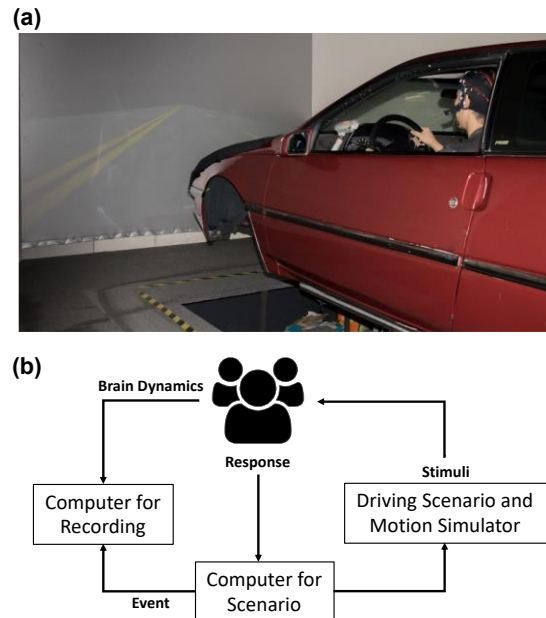


Figure 1: Experimental environment: (a) a 360-degree virtual-reality lab with a motion platform in BRC, NCTU. (b) The experimental setup: the computer used for recording can synchronize all response and brain dynamics from the participants and the relevant events from the experimental scenario and thus detects the responses of the participants.

virtual-reality laboratory with a motion platform using a response-detecting steering wheel (Fig. 1), which simulates the real-world driving environment [11], [28]. The vehicle in the experiment was set at 100 km/h on a four-lane road at night. The vehicle deviated randomly from the position at the center of the cruising lane (3rd lane, as shown in Fig. 2). The participants were asked to drive the vehicle back to the center of the lane using the steel wheel as soon as possible when the participant became aware of the deviation. There was one resting period between two continuous trials, and this period was one random number from 7 to 12 seconds to avoid

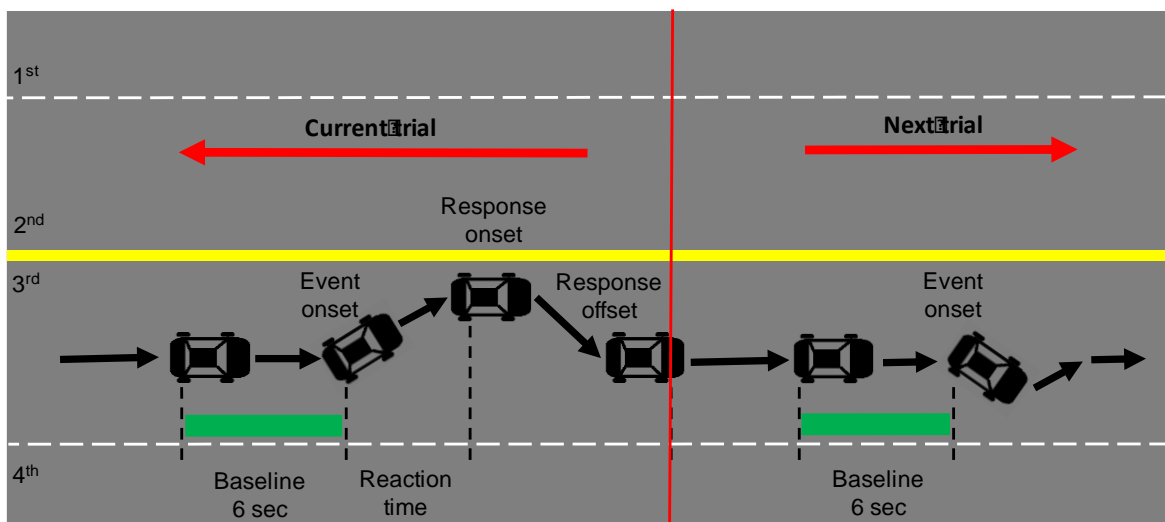


Figure 2: Experimental scenario in this study. The cruising car randomly drifted to the left (2nd) or right (4th) lane, and the participants were asked to control the car back to the cruising (3rd) lane. We used 6-second EEG data before the event onset (baseline, marked by bold green lines) as the input data for 4D CNN, and the period between the event onset and the response onset was defined as the response time. The period between two continuous trials was approximately 7 to 12 seconds.

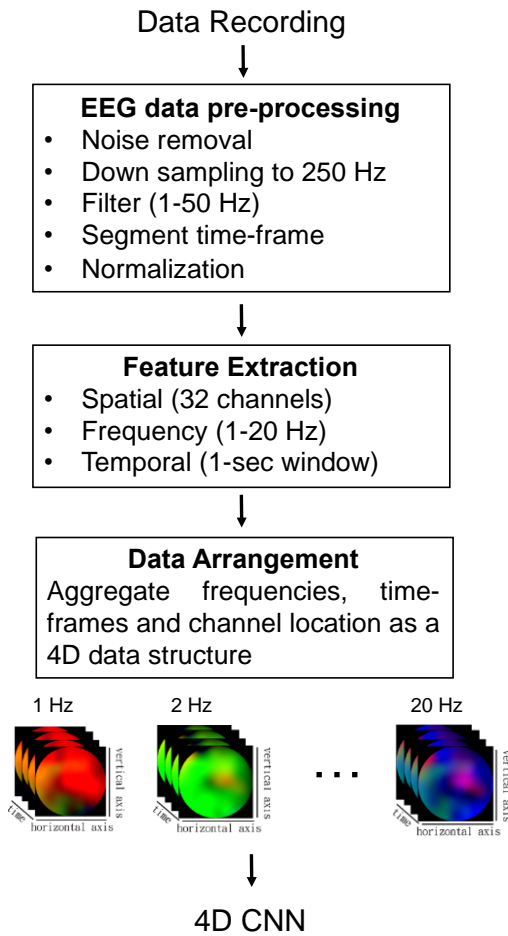


Figure 3: Structure of data pre-processing and feature extraction in the 4D CNN algorithm. For each trial, the re-sampling, bandpass filter and noise removal were adopted to the 6-second EEG baseline. FFT was then applied to translate the frequency information. Each of the 2D topo-images included the frequency power of 30 EEG channels in a single timeframe.

participants' expectation of coming deviation. The RT was measured to evaluate the driver's driving performance in the study. RT is the latency between the onset of the deviation and the onset of the response (as shown in Fig. 2). A high RT indicates that the participant is relatively drowsy in this VR-highway driving scenario, and a low RT indicates that the participant that is relatively alert.

The driving experiment was conducted in the early afternoon (13:00–15:00) after lunch, when the circadian rhythm of drowsiness is at its peak [29]. In addition, the VR-highway scene was monotonous, and the task demand was low to induce drowsiness [11], [30], [31]. Under such conditions, the subjects had difficulty regulating attention and performance, which resulted in long RTs [32]. If the participants did not properly react to the designed tasks during the alert situation, the related data were excluded from the subsequent analyses.

The entire dataset included 71 sessions from 37 healthy participants, and there were 60 EEG epochs in each session. The participants did not have any history of psychological or sleep disorders. All EEG data were collected through a 32-channel EEG cap with Neuroscan system, including 30 regular channels and two reference channels. The impedances of the electrodes were all less than 5 k Ω , and the sampling rate of the EEG

signals was 500 Hz. To ensure that the brain state of the participants would not be affected by external factors, consumption of alcohol or caffeinated drinks and strenuous exercise were prohibited for the one week prior to the experiments. All participants were clearly informed about the operation of the related equipment and the procedure of the experiment and had practiced the procedure sufficiently before participating in the experiment. Furthermore, before the experiment, all of the subjects were asked to read and sign an informed consent form. Finally, this experiment was approved by the Institutional Review Board of the Veterans General Hospital, Taipei, Taiwan.

B. Data processing

Fig. 3 shows the diagram for processing EEG signals including three major components, EEG data pre-processing, feature extraction, and Data Arrangement for 4D CNN.

EEG data pre-processing

We applied EEGLAB for EEG data pre-processing [39]. EEGLAB is an open source Matlab toolbox and widely applied to EEG data processing [11], [29], [34], [36], [39]. The noises associated with eye movements and muscle activities were removed manually, followed by 250-Hz downsampling, a high-pass filter (1 Hz) and a low-pass filter (50 Hz). The EEG signals were later segmented into a 6-second baseline signal, which began 6 seconds before event onset [11], [29]–[31]. To eliminate inter-subject differences, the measured RTs and EEG dynamics were normalized by the trials with 10% fastest RTs in every single subject [11], [12]. The datasets collected from each individual subject were then normalized by subtracting the mean and dividing by the standard deviation.

Feature Extraction

The EEG signals were then transformed into a 4D format that included a frequency power of 1-20 Hz, 6-second temporal information and a two-dimension topo-image. The bottom of Fig. 3 shows the data structure of our 4D input data, and each topo-image was converted from the frequency power of all 30 EEG channels in one single timeframe using the linear interpolation function of EEGLAB [39], one of the most powerful EEG analysis toolboxes in MATLAB. This interpolation is able to map the scalp data into a 2D view, and each topology image depicts the power distributions for one specific frequency and one single timeframe. The more details about temporal information, power distributions and channel locations are as the following paragraphs.

The major advantage of EEG signals is high temporal resolution. To capture cognitive changes within consecutive time points, temporal convolution was introduced in this study. A previous study [33] demonstrated that motion patterns on video were successfully extracted by temporal convolution in the consecutive images. Following this path, 4D CNN uses data from a 1-second time window as one input frame, and the overlap between two input frames was 600 milliseconds.

In previous EEG studies [32], [34]–[37], frequency oscillations in the delta (1-4 Hz) theta (4-7 Hz), alpha (8-13 Hz) and low beta (13-20 Hz) bands were significant indicators for monitoring cognitive performance. Especially, the power oscillations in theta and alpha bands were highly associated

Table 1: Configuration of 4D CNN, 3D CNN, CNN, ANN and LSTM

	4D CNN	3D CNN	CNN	ANN, LSTM
Input layer				
1st convolutional layer	16 kernels at size 3×3×3×3	16 kernels at size 3×3×3	16 kernels at size 3×3	128 nodes
1st max-pooling layer	kernel size 2×2×2×2	kernel size 2×2×2	kernel size 2×2	
2nd convolutional layer	32 kernels at size 3×3×3×3	32 kernels at size 3×3×3	32 kernels at size 3×3	
2nd pooling layer	kernel size 2×2×2×2	kernel size 2×2×2	kernel size 2×2	
3rd convolutional layer	64 kernels at size 3×3×3×3	64 kernels at size 3×3×3	64 kernels at size 3×3	
3rd pooling layer	kernel size 2×2×1×1	kernel size 2×1×1	kernel size 2×1	
Fully connected layer	128 nodes	128 nodes	128 nodes	
Output layer				

with sustained attention [17]–[20]. However, most EEG studies observe the data orientation by averaging in one frequency band. There is a high possibility of ignoring essential information due to the elimination of fluctuations in frequency bands. To explore hidden information in frequency domains, this study considered the frequencies between 1 and 20 Hz. Fast Fourier Transform (FFT) was adopted.

Data Arrangement for 4D CNN

Fig. 3 shows the 4D data format of input data. The human brain is dominated by four major lobes, the frontal, parietal, temporal and occipital lobes. Each lobe is responsible for distinct cognitive functions [9], [31]–[34]. Electrodes are placed around the scalp to acquire brain information, and a 32-channel EEG cap was used in this study. In a previous study [22], the approach of using the topographic map (topo-map) was shown to be effective for tracking brain activity, which implies that the signals are inseparably associated with the electrode positions. To include channel location in the input data, a 2D topology image with a size of 32×32. For each trial, the brain dynamics 6 seconds before the deviation were extracted as the baseline signals. The extracted baseline signals were segmented into overlapping 1000-ms samples that were advanced in 600-ms steps, which totally resulted in 13 samples from each 6000-ms baseline. For each sample, 20 frequency bins (1 Hz - 20 Hz) were extracted by FFT. The feature dimension of each 1000-ms sample is 20 (20 frequency bins). By organizing input data into a 4D format, we were able to access the hidden information in the three domains (spatial, frequency and temporal domains) with a size of 32*32*20*13.

The strength of learning visual patterns from local neighboring area (pixel) saves us from designing hand-craft feature extractors [23]. A typical CNN structure is composed of convolutional layers, pooling layers and fully connected layers. Along with dataflow, the low-level features (edges, lines, corners) are combined, evolved into high-level features and form abstract interpretations.

C. 4D CNN

This study proposes a 4D CNN for forecasting driving performance. 4D CNN is a CNN-based deep learning algorithm that is designed for the analysis of EEG signals, which includes more data dimensions than 2D and 3D CNNs. At the convolutional layer, 4D convolution is performed and defined as follows:

$$v_{ij}^{wxyz} = \text{relu} \left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} \sum_{s=0}^{S_i-1} W_{ijm}^{pqrs} v_{(i-1)m}^{(w+p)(x+q)(y+r)(z+s)} \right)$$

where v_{ij}^{wxyz} is the unit of the w , x , y and z positions of the feature map between the i th and j th layers; b_{ij} is the bias between the two layers; m represents the m^{th} filter in the convolutional layer; P , Q , R and S are the kernel size of the filters; W_{ijm}^{pqrs} are the kernel weights at the pqr position of the m^{th} filter; and relu is an activation function for accelerating the convergence speed during training [23].

Each unit in the output feature map is the result of the inner product of kernels and feature map of the previous layer, where $v_{(i-1)m}^{(w+p)(x+q)(y+r)(z+s)}$ is the feature map of the previous layer. Bias and kernel weights are trained by backpropagation, following the same approach as in CNN. In the pooling layer, a 4D pooling layer is adopted to down-sample the data points of the feature maps.

In the current study, the 4D CNN net was designed as nine layers including an input layer, a convolutional layer (16 filters), a max-pooling layer, a convolutional layer (32 filters), a max-pooling layer, a convolutional layer (64 filters), a max-pooling layer, a fully connected layer (128 nodes) and an output layer. All of the activation functions in the convolutional layers and the fully connected layer were applied using rectified linear units (ReLU).

The 4D CNN algorithm was trained through backpropagation with a batch size of eight. The optimizer was Adaptive Moment Estimation (ADAM). The initial learning

Table 2: Results of 4D CNN, 3D CNN, CNN, ANN, LSTM and SVR

Model	RMSE		Error rate		Correlation coefficient		Number of trainable parameters
	Average	Improvement	Average	Improvement	Average	Improvement	
ANN	0.619±0.137	8.89 %	0.316±0.047	10.40 %	0.210±0.186	31.65%	76,928
CNN	0.599±0.149	5.91 %	0.297±0.061	4.88 %	0.217±0.171	27.47%	121,857
LSTM	0.616±0.145	8.50 %	0.318±0.045	10.96 %	0.106±0.082	161%	373,377
SVR	0.620±0.249	9.00 %	0.317±0.047	10.74 %	0.216±0.135	28.11%	—
3D CNN	0.586±0.136	3.82 %	0.293±0.052	3.45 %	0.247±0.187	11.98%	1,544,481
4D CNN	0.564±0.139	—	0.283±0.066	—	0.277±0.182	—	2,830,465

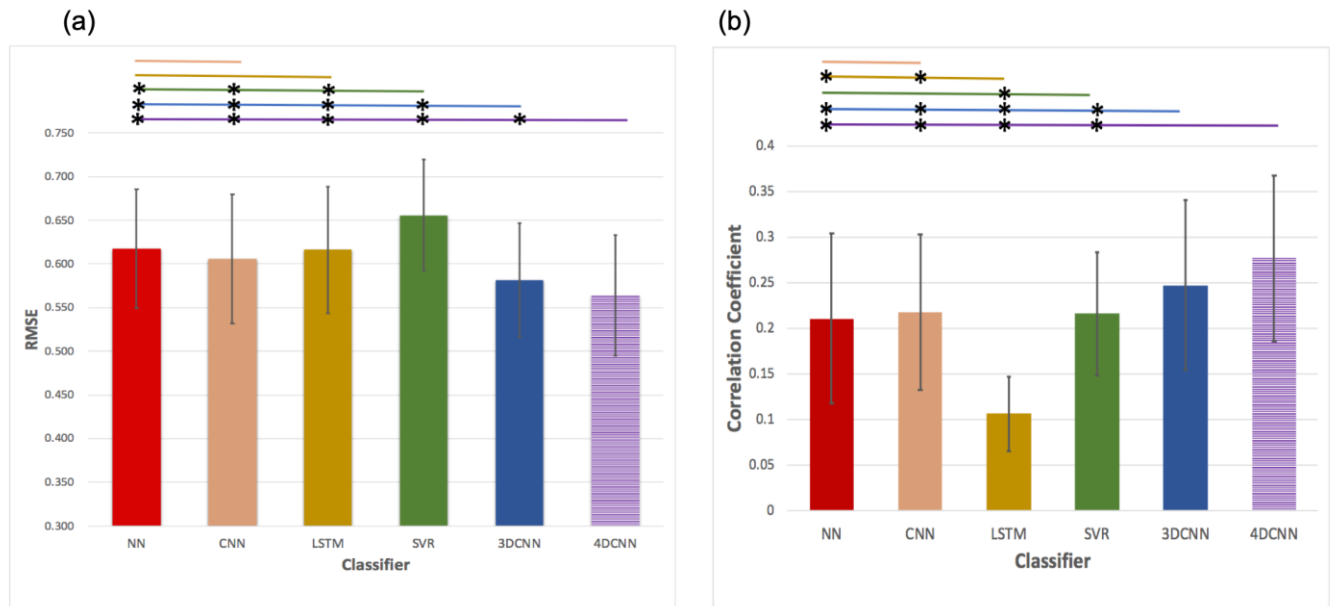


Figure 4: Performance of (a) RMSE (b) correlation coefficient among all models. Star indicates p-value is less than 0.05.

rate was 0.001 and was divided by five when the validating error improvement (10% of the training data) was less than 1%.

D. Evaluation Criteria

This study compared the system performance with state-of-the-art algorithms, namely, 3D CNN, CNN, artificial neural network (ANN), support vector regression (SVR) and Long Short-Term Memory (LSTM). Table 1 lists the structures and parameters of 4D CNN, 3D CNN, CNN, ANN and LSTM. To ensure comparability, the network configurations for 4D CNN, 3D CNN and CNN are identical, e.g., the width of the convolutional layer is 1. The only difference is the kernel size for fitting with input data. For input data, the 4D CNN algorithm uses a 4D topo-image of size $32 \times 32 \times 20 \times 13$ (x-axis and y-axis of the topo-image, timeframes and frequency bins). 3D CNN and LSTM takes data of size $32 \times 13 \times 20$ (EEG channels, timeframes and frequency bins). CNN, ANN and

SVR takes in 2D images of size 32×20 (EEG channels and frequency).

In the experiments, leave-one-subject-out cross validation was used. Seventy sections (4200 epochs) were taken as the training set, 10% of the training set was used as the validation set (420 epochs), and one section was used as the testing set (60 epochs). These steps were repeated 71 times to complete the cross validation (71 experimental sessions). To prevent overfitting, this study adopted an early stop: the training was terminated if there was no significant improvement in the validation performance for consecutive 700 iterations.

The root mean square error (RMSE), correlation coefficient and the error ratio between the predicted driving performance and the ground truth (the observed driving performance) were introduced to evaluate the performance. The RMSE represented the standard deviation between the predicted driving performance and ground truth. The error ratio and correlation

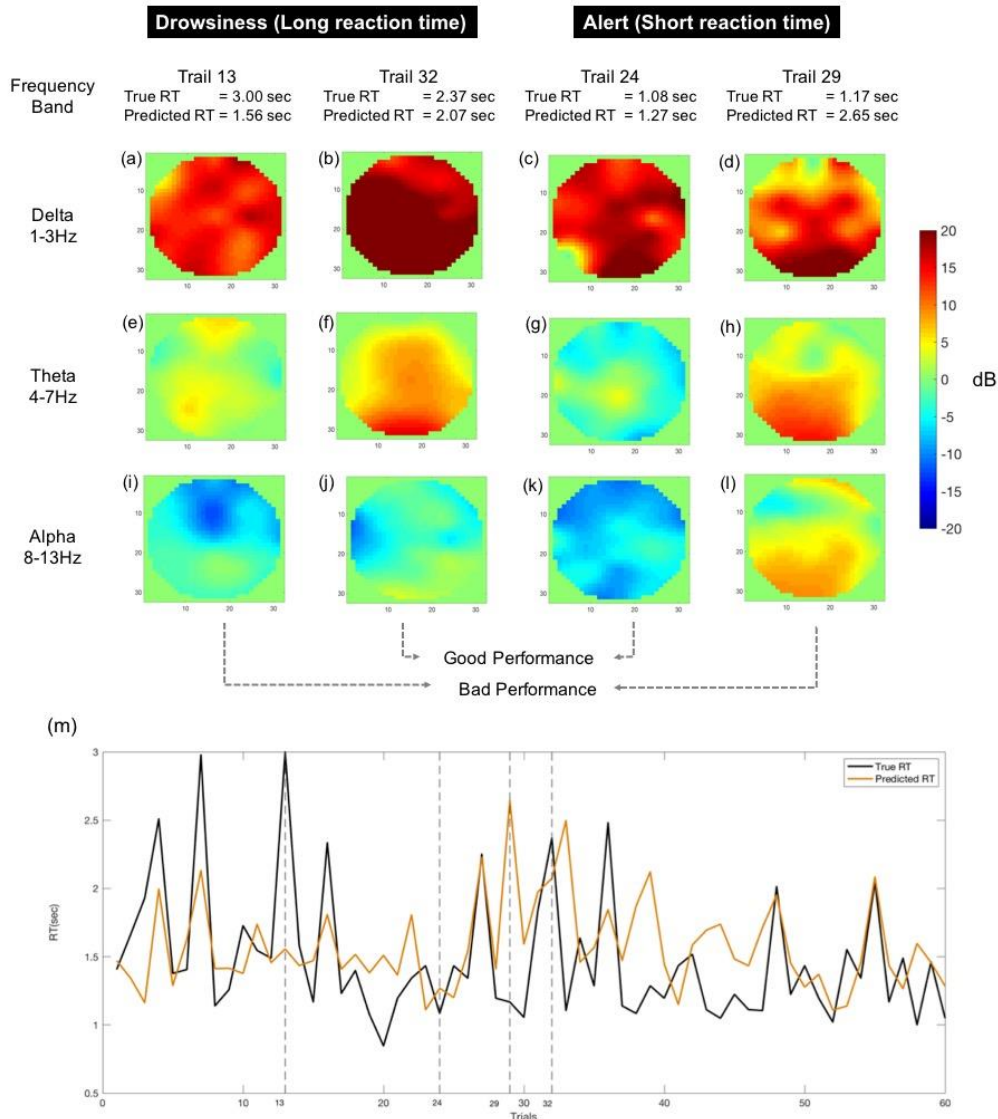


Figure 5: The topo-images of brain activity. (a)-(d) show the topo-images of delta band for four trails. (e)-(h) show the topo-images of theta band for four trails. (i)-(l) show the topo-images of alpha band for four trails. (m) A comparison of the predicted RT of 4D CNN (orange line) and the true RT (black line) in a single session.

coefficient were used to evaluate the overall performance, which is calculated as follows:

$$\frac{|observed\ driving\ performance - predicted\ driving\ performance|}{observed\ driving\ performance}$$

4D CNN was programmed using Python 2.7 and TensorFlow 1.0 and was run on a cluster with 128 GB RAM, 2400 MHz, DDR4, Intel Xeon E5-2680 v3 CPU and NVIDIA Quadro K2200 GPU to boost the training speed.

III. RESULTS AND DISCUSSION

A. Prediction Performance

Table 2 shows the predicted driving performance obtained with 4D CNN, 3D CNN, CNN, ANN, SVR and LSTM. The first two columns show the RMSE values and improvement rates. The RMSE of 4D CNN was 0.5638 ± 0.1391 , which

corresponds to an improvement in performance of 3.82%, 5.91%, 8.89%, 9.00% and 28.11% compared with 3D CNN, CNN, ANN, SVR and LSTM, respectively. The error rate of 4D CNN also outperformed the error rates of the other algorithms: the error rate of 4D CNN was 3.45% better than that of 3D CNN and 4.88% better than that of CNN. Similarly, 4D CNN showed 10.4% and 10.74% improved error rates compared with those of the shallow algorithms ANN and SVR, respectively, and a 10.96% improvement in the error rate compared with LSTM. 4D CNN also showed a relatively high correlation coefficient among the other models.

To evaluate whether the improvement was statistically significant, the RMSE and correlation coefficient were evaluated using a T-test, as shown in Fig. 4. A star indicates a p-value less than 0.05. As shown in Fig. 4(a), the RMSE performance of 4D CNN was significantly different compared with that of the other models. The same finding was obtained

for the correlation coefficients of the other models, with the exception of 3D CNN, as illustrated in Fig. 4(b).

B. Changes in the Human Brain

It is worth investigating whether the features learned by 4D CNN can effectively predict the fatigue level. Thus, we focused on exploring the relationships between the changes in driving performance and human brain dynamics. In the following discussion, we show the topo-images in the delta, theta and alpha bands under alert and drowsy conditions.

Fig. 5(m) depicts the true driving performance (black line) and predicted driving performance by 4D CNN (orange line) in one session. To investigate the brain dynamics under good driving performance (shorter RTs) and bad driving performance (long RTs), four EEG epochs were selected for further discussions. Epochs 13 and 24 were selected due to the short real RTs. Epochs 29 and 32 listed longer real RTs. The predicted RTs and the real RTs for these four selected EEG epochs are listed in Fig. 5. Moreover, there were two EEG epochs (epoch 13 and epoch 29) with large errors (bad performance) and two trials (epoch 24 and epoch 32) with small predicted errors (good performance). The participant might not feel drowsy while responding to the car deviation in EEG epoch 13 since there were low real RTs in the previous and following EEG epochs. Therefore, the participant may still be at a high arousal level, and the low RT was predicted based on the recorded EEG signals. In EEG epoch 29, a large difference between the real RT and predicted RT was discovered. The trend in real RTs was increased from epoch 24, and the increased trend indicates that the participant might start to feel drowsy. However, the participant may still handle the lane-keeping task with accepted performance as he/she notices the deviation in the beginning of EEG epoch. One can have similar experiences in real life; for example, one can still maintain the car to avoid accidents despite feeling tired. In this case, low real RT and brain patterns with a low level of arousal should be recorded. In this study, the RTs were predicted based on the brain dynamics and the 4D CNN model. The high predicted RT can be imaged in this trained 4D CNN model since the recorded brain patterns in EEG epoch 29 might indicate that the participant already feels sleepy.

Brain Dynamics

Fig. 5(a)-(l) depict the human brain changes in the selected EEG epochs with long and short RTs to explore the relationships between brain dynamics in specific bands and driving performance. Fig. 5(a)-(d), (e)-(h), and (i)-(l) show the topo-images of the delta (1-3 Hz), theta (4-7 Hz) and alpha (8-13 Hz) bands, respectively. The subfigures in the second and third rows were extracted from the EEG epoch with a small difference between the real RTs and predicted RTs. On the other hand, the subfigures in the first and fourth rows list the brain patterns in the EEG epochs with large differences between the predicted RTs and real RTs.

According to previous studies, the brain dynamics in the theta and alpha bands are positively correlated with the attentional demands or levels of arousal [31], [38]. In particular, researchers have mainly focused on the changes in the frontal and posterior cingulate cortex (parietal and occipital lobes) [36]-[38]. The longer predicted RTs in epochs 32 (2.37 seconds)

and 29 (2.65 seconds) show that the brain patterns can be linked to the related phenomenon at a low level of arousal. In EEG epochs 32 (Fig. 5(f) and (j)) and 29 (Fig. 5(h) and (l)), the higher posterior theta, frontal alpha, and posterior alpha can also be observed. On the other hand, the brain dynamics in epochs 13 (1.56 seconds) and 24 (1.27 seconds) are associated with a high level of arousal due to the shorter predicted RTs. The decreased posterior theta and alpha activations can also be extracted through the proposed 4D CNN model (ref. Fig. 5(e), (i), (g) and (k)). In short, the proposed 4D CNN obtained a higher RT while there were increased theta and alpha activations in the posterior region. These results in Table 2 and Fig. 5 show that 4D CNN has better ability to extract the significant brain features for forecasting driving performance.

Deep learning is a powerful tool which can be applied to various fields. In the current research, the EEG patterns with spatial (32 channels, two dimensions), frequency (1-20 Hz, one dimension) and temporal (13 frames, one dimension) information have been used. As the results are shown in Table 2, the performance is the best but more trainable parameters are also required. The time cost for EEG feature processing and driving performance prediction is associated with capabilities of hardware. Previous researches indicated the normal driving performance (without drowsiness) was in the range of 0.6 – 1.2 second [31], [34]-[36]. In this experiment, it takes 0.1441 ± 0.0036 second for EEG feature processing and 0.2197 ± 0.0075 second for driving performance prediction on the well-trained 4DCNN model which is less than the real driving performance in the awake situation. There is no doubt that the processing time of applying 4DCNN model on driving performance estimation can be further improved with state-of-the-art technology such as cloud, fog and parallel computing. Moreover, the size of data set is relatively small. First, the size of data set is relatively small. There are totally 71 experimental sessions from 37 participants in the current research. The subjective difference is always a big problem in modelling biosignals and brain-computer interface, especially the real-life application. We will conduct large scale experiment to collect more physiological information in a long-term period. Second, more cognitive tasks can be included to simulate the real driving environment in which the changes of cognitive workload, stress or fatigue can be assessed for building more stable models. Finally, the 71 experimental sessions are collected in one VR-based driving scenario. However, the real-world applications are the goals in near future. Therefore, some real-world experiments will be conducted to test the performance and stability as getting ethics approval from Institutional Review Board.

IV. CONCLUSION

This study proposes a 4D CNN-based algorithm to forecast driver performance based on recorded EEG signals. This 4D CNN approach helps narrow down the position of significant features and preserve the signal trend for monitoring purposes. To demonstrate the capability of 4D CNN, three indexes (RMSE, error rate and correlation coefficient) were evaluated for comparison purposes. The RMSE of 4D CNN was improved compared with those of 3D CNN, CNN, ANN, SVR and LSTM by 3.82%, 5.91%, 8.89%, 9.00% and 8.5%, respectively.

Meanwhile, the error rate of 4D CNN was better than the error rates of 3D CNN, CNN, ANN, SVR and LSTM by 3.45%, 4.88%, 10.4%, 10.74% and 10.96%, respectively. The analysis of the correlation coefficients showed the competitive performance of 4D CNN compared with the other models. The P-values of the RMSE and correlation coefficient demonstrated the significant improvement of the application of 4DCNN over the other methods. The relationships between the changes in driving performance and human brain dynamics were also explored through the characteristics of the CNN. These results indicate that 4D CNN can extract essential features from high-dimensional EEG data to accurately forecast the fatigue level. Based on the final performance and these significant features, we envision that deep learning might open a new branch between translation neuroscience and real-world applications, but the current research is far from real-life application. In further research, we will leverage the pretrained model and image processing technologies to build a closed-loop system for monitoring the fatigue level and detecting the onset of lane deviation, respectively. This new research might benefit real-world applications (e.g., autonomous driving) in the near future.

ACKNOWLEDGMENTS

This work was supported in part by the Australian Research Council (ARC) under discovery grant DP180100670 and DP180100656. Research was also sponsored in part by the Australia Defence Innovation Hub under Contract No. P18-650825 and US Office of Naval Research Global under Cooperative Agreement Number ONRG - NICOP - N62909-19-1-2058. We also thank the NSW Defence Innovation Network and NSW State Government of Australia for financial support in part of this research through grant DINPP2019 S1-03/09. The work was also supported by Taiwan Ministry of Science and Technology under the Grant MOST 108-2321-B-038-005-MY2 and MOST108-2628-E-019-001-MY3.

REFERENCES

- [1] World Health Organization. Dept. of Violence, et al. "Global Status Report on Road Safety: Time for Action," *World Health Organization*, 2009.
- [2] T. Rosenbloom, A. Shahar, A. Elharar, and O. Danino, "Risk perception of driving as a function of advanced training aimed at recognizing and handling risks in demanding driving situations," *Accid. Anal. Prev.*, vol. 40, no. 2, pp. 697–703, 2008.
- [3] World Health Organization, "World Health Statistics 2008," *World Health Organization*, 2008.
- [4] I. D. Brown, "Driver fatigue," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 36, pp. 298-314, 1994.
- [5] P. Thiffault and J. Bergeron, "Monotony of road environment and driver fatigue: a simulator study," *Accid. Anal. Prev.*, vol. 35, no. 3, pp. 381–391, 2003.
- [6] Y. I. Noy, W. J. Horrey, S. M. Popkin, S. Folkard, H. D. Howarth, and T. K. Courtney, "Future directions in fatigue and safety research," *Accid. Anal. Prev.*, vol. 43, no. 2, pp. 495–497, 2011.
- [7] K. Idogawa, "On the brain wave activity of professional drivers during monotonous work," *Behaviormetrika*, vol. 18, pp. 23-34, 1991
- [8] T. J. Balkin, W. J. Horrey, R. C. Graeber, C. A. Czeisler, and D. F. Dinges, "The challenges and opportunities of technological approaches to fatigue management," *Accid. Anal. Prev.*, vol. 43, no. 2, pp. 565–572, 2011.
- [9] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neurosci. Biobehav. Rev.*, vol. 44, pp. 58–75, 2014.
- [10] C. Zhao, C. Zheng, M. Zhao, Y. Tu and J. Liu, "Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic," *Expert Systems with Applications*, 2010
- [11] R. Wu, C. Lin and S. Liang, "Estimating driving performance based on EEG spectrum and fuzzy neural network", *2004 IEEE International Joint Conference on Neural Networks*, 2004.
- [12] C. T. Lin et al., "Wireless and Wearable EEG System for Evaluating Driver Vigilance," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 2, pp. 165–176, 2014.
- [13] R. R. Johnson, D. P. Popovic, R. E. Olmstead, M. Stikic, D. J. Levendowski, and C. Berka, "Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model," *Biol. Psychol.*, vol. 87, no. 2, pp. 241–250, 2011.
- [14] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver Drowsiness Classification Using Fuzzy Wavelet-Packet-Based Feature-Extraction Algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, 2011.
- [15] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver Inattention Monitoring System for Intelligent Vehicles: A Review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, 2011.
- [16] R. Rosipal et al., "EEG-Based Drivers' Drowsiness Monitoring Using a Hierarchical Gaussian Mixture Model," in *International Conference on Foundations of Augmented Cognition*, pp. 294–303, 2007.
- [17] A. Sharma, "Epileptic seizure prediction using power analysis in beta band of EEG signals," in *2015 International Conference on Soft Computing Techniques and Implementations (ICSCITI)*, pp. 117–121, 2015.
- [18] O. Aydemir, "Olfactory Recognition Based on EEG Gamma-Band Activity," *Neural Comput.*, vol. 29, no. 6, pp. 1667–1680, 2017.
- [19] C. Dissanayaka et al., "Sleep onset detection with multiple EEG alpha-band features: Comparison between healthy, insomniac and schizophrenic patients," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp. 1–4, 2015.
- [20] A. Chatchinarat, K. W. Wong, and C. C. Fung, "A comparison study on the relationship between the selection of EEG electrode channels and frequency bands used in classification for emotion recognition," in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 1, pp. 251–256, 2016.
- [21] S. Terachi and Y. Tanaka, "Detailed study of temporal variations of EEG by improved MEM spectral analysis with high resolution power," in *Proceedings of the Sixth International Symposium on Signal Processing and its Applications*, vol. 1, pp. 230–233 vol.1, 2001
- [22] L. S. Hooi, H. Nisar, and Y. V. Voon, "Comparison of motion field of EEG topo-maps for tracking brain activation," in *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 251–256, 2016.
- [23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [24] O. Abdel-Hamid, A. r Mohamed, H. Jiang, and G. Penn, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4277–4280, 2012.
- [25] L. Jingwei, C. Yin, and Z. Weidong, "Deep learning EEG response representation for brain computer interface," in *2015 34th Chinese Control Conference (CCC)*, pp. 3518–3523, 2015.
- [26] M. Hajinoroozi, Z. Mao, and Y. Huang, "Prediction of driver's drowsy and alert states from EEG signals with deep learning," in *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 493–496, 2015.
- [27] N. Lu, T. Li, X. Ren, and H. Miao, "A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, 2017.
- [28] S. Stober, D. J. Cameron, and J. A. Grahn, "Using Convolutional Neural Networks to Recognize Rhythm Stimuli from Electroencephalography Recordings," in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Cambridge, MA, USA, pp. 1449–1457, 2014.
- [29] C. T. Lin, I. F. Chung, L. W. Ko, Y. C. Chen, S. F. Liang and J. R. Duann, "EEG-Based Assessment of Driver Cognitive Responses in a Dynamic Virtual-Reality Driving Environment", *IEEE Trans. Biomed.*

- Eng, 2007.
- [30] M. Ferrara and L. D. Gennaro, "How much sleep do we need?," *Sleep Med. Rev.*, vol. 5, no. 2, pp. 155–179, 2001.
- [31] K. C. Huang, C. H. Chuang, Y. K. Wang, C. Y. Hsieh, J. T. King and C. T. Lin, "The effects of different fatigue levels on brain-behavior relationships in driving," *Brain and Behavior*, vol. 9, no. 12, 2019.
- [32] C. T. Lin, J. T. King, C. H. Chuang, W. Ding, W. Y. Chuang, L. D. Liao, and Y. K. Wang, "Exploring the brain responses to driving fatigue through simultaneous EEG and fNIRS measurements," *International Journal of Neural Systems*, vol. 30, 2020.
- [33] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional Neural Networks for Human Action Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, 2013.
- [34] Y. K. Wang, T. P. Jung, and C. T. Lin, "EEG-based attention tracking during distracted driving," *IEEE Trans. Neural Systems and Rehabilitation Engineering*, vol. 23, no.6, pp. 1085–1094, 2015.
- [35] C. T. Lin, C. H. Chuang, Y. K. Wang, S. F. Tsai, T. C. Chiu and L. W. Ko, "Neurocognitive characteristics of the driver: A review on drowsiness, distraction, navigation, and motion sickness," *Journal of Neuroscience and Neuroengineering*, 1(1), pp. 61-81, 2012.
- [36] K. C. Huang, T. Y. Huang, C. H. Chuang, J. T. King, Y. K. Wang, C. T. Lin, and T. P. Jung, "An EEG-based fatigue detection and mitigation system," *International journal of neural systems*, 26(04), 1650018, 2016.
- [37] C. H. Chuang, Z. Cao, J.T. King, B.S. Wu, Y. K. Wang and C.T. Lin, "Brain Electrodynamical and Hemodynamic Signatures Against Fatigue During Driving," *Frontiers in Neuroscience*, 12, pp. 181., 2018
- [38] Y. K. Wang, T. P. Jung, and C. T. Lin. "Theta and alpha oscillations in attentional interaction during distracted driving," *Frontiers in behavioral neuroscience* 12, pp. 3, 2018
- [39] A. Delorme and S. Makeig. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods*, 134(1): 9-21, 2014.
- [40] R. Hefron, B. Borghetti, C. S. Kabban, J. Christensen, and J. Estépp, "Cross-Participant EEG-Based Assessment of Cognitive Workload Using Multi-Path Convolutional Recurrent Neural Networks," *Sensors (Basel, Switzerland)*, 18(5), 2018.
- [41] R. G. Hefron, B. J. Borghetti, J. C. Christensen, and C. M. S. Kabban, "Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation," *Pattern Recognition Letters*, 94, pp. 96–104, 2017.
- [42] R. T. Schirmer, J. T. Springenberg, L.D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, 38(11), pp. 5391–5420, 2017.
- [43] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, 15(5), 2018.
- [44] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, 16(3), 2019.
- [45] Y. Ming, W. Ding, D. Pelusi, D. Wu, Y. K. Wang, M. Prasad, and C. T. Lin, "Subject adaptation network for EEG data analysis," *Applied Soft Computing*, 80, 2019.
- [46] Y. Ming, D. Pelusi, C. N. Fang, M. Prasad, Y. K. Wang, D. Wu, and C. T. Lin, "EEG data analysis with stacked differentiable neural computers," *Neural Computing and Applications*, 32(12), pp. 7611-7621, 2020.

Chin-Teng Lin (S'88–M'91– SM'99–F'05) received the B.S. degree from National Chiao-Tung University (NCTU), Taiwan in 1986, and the Master and Ph.D. degree in electrical engineering from Purdue University, USA in 1989 and 1992, respectively. He is currently Distinguished Professor of Faculty of Engineering and Information Technology, and Co-Director of Center for Artificial Intelligence, University of Technology Sydney, Australia. He is also invited as Honorary Chair Professor of Electrical and Computer Engineering, NCTU, International Faculty of University of California at San-Diego (UCSD), and Honorary Professorship of University of Nottingham. Dr. Lin was elevated to be an IEEE Fellow for his contributions to biologically inspired information systems in 2005, and was elevated International Fuzzy Systems Association (IFSA) Fellow in 2012. Dr. Lin received the IEEE Fuzzy Systems Pioneer Awards in 2017. He served as the Editor-in-chief of IEEE Transactions on

Fuzzy Systems from 2011 to 2016. He also served on the Board of Governors at IEEE Circuits and Systems (CAS) Society in 2005-2008, IEEE Systems, Man, Cybernetics (SMC) Society in 2003-2005, IEEE Computational Intelligence Society in 2008-2010, and Chair of IEEE Taipei Section in 2009-2010. Dr. Lin was the Distinguished Lecturer of IEEE CAS Society from 2003 to 2005 and CIS Society from 2015-2017. He serves as the Chair of IEEE CIS Distinguished Lecturer Program Committee in 2018-2019. He served as the Deputy Editor-in-Chief of IEEE Transactions on Circuits and Systems-II in 2006-2008. Dr. Lin was the Program Chair of IEEE International Conference on Systems, Man, and Cybernetics in 2005 and General Chair of 2011 IEEE International Conference on Fuzzy Systems. Dr. Lin is the coauthor of *Neural Fuzzy Systems* (Prentice-Hall), and the author of *Neural Fuzzy Control Systems with Structure and Parameter Learning* (World Scientific). He has published more than 330 journal papers (Total Citation: 22,778, H-index: 67, i10-index: 288) in the areas of neural networks, fuzzy systems, brain computer interface, multimedia information processing, and cognitive neuro-engineering, including about 125 IEEE journal papers.

Chun-Hsiang Chuang received the B.S. degree in mathematics education from Taipei Municipal Teachers College, Taipei, Taiwan, in 2004, the M.S. degree in educational measurement and statistics from National Taichung University, Taichung, Taiwan, in 2009, and the Ph.D. degree in electrical engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2014. He was a Visiting Scholar with the Swartz Center for Computational Neuroscience, University of California at San Diego, La Jolla, CA, USA, from 2012 to 2013. He is currently an Assistant Professor with the National Tsing Hua University, Hsinchu, Taiwan. His current research interests include machine learning, computational neuroscience, bio-medical signal processing, and brain-computer interface.

Yu-Chia Hung received the B.S. from National Sun Yat-sen University, Kaohsiung, Taiwan 2014 and M.S. degree from National Chiao Tung University, Hsinchu, Taiwan in 2016. His current research interests include brain computer interface, deep learning and signal processing.

Chieh-Ning Fang received the B.S. degree from the Department of Electrical and Computer Engineering in 2014 and M.S. degree from Institute of Electrical and Control Engineering in 2016, National Chiao-Tung University, Hsinchu, Taiwan. Her research interests cover deep learning, neural network, image recognition and pattern recognition. Future research aims at developing a brain computer interface associated with deep learning algorithm to improve its performance.

Dongrui Wu (S'05–M'09–SM'14) received the BE degree in automatic control from the University of Science and Technology of China in 2003, the ME degree in electrical engineering from the National University of Singapore in 2005, and the PhD degree in electrical engineering from the University of Southern California in 2009. He is now Professor in the School of Automation, Huazhong University of Science and Technology, Wuhan, China. His research interests include affective computing, brain-computer interface, computational intelligence, and machine learning. He has more than 100 publications, including a book entitled *Perceptual Computing* (Wiley-IEEE, 2010). Dr. Wu received the IEEE International Conference on Fuzzy Systems Best Student Paper Award in 2005, the IEEE CIS Outstanding PhD Dissertation Award in 2012, the IEEE TRANSACTIONS ON FUZZY SYSTEMS Outstanding Paper Award in 2014, the NAFIPS Early Career Award in 2014, and the IEEE SMC Society Early Career Award in 2017. He was also a finalist of another three Best Paper Awards. He is an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, the IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, and the IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE.

Yu-Kai Wang (M'13) received the B.S. degree in mathematics education from National Taichung University of Education, Taichung, Taiwan, in 2006, the M.S. degree in biomedical engineering from National Chiao Tung University (NCTU), Hsinchu Taiwan, in 2009, and the Ph.D. degree in computer science from NCTU, Hsinchu Taiwan, in 2015. He is currently a Lecturer of Faculty of Engineering and Information Technology, and Co-Director of Motion Platform and Mixed Reality Lab, University of Technology Sydney, Australia. His current research interests include computational neuroscience, human performance modelling, biomedical signal processing, and brain-computer interface.