Exploring Covert States of Brain Dynamics via Fuzzy Inference Encoding

Yu-Cheng Chang, Yu-Kai Wang, *Member, IEEE*, Nikhil R. Pal, *Fellow, IEEE*, and Chin-Teng Lin, *Fellow, IEEE*

***Abstract*— Human brain inherently exhibits latent mental processes which are likely to change rapidly over time. A framework that adopts a fuzzy inference system is proposed to model the dynamics of the human brain. The fuzzy inference system is used to encode real-world data to represent the salient features of the EEG signals. Then, an unsupervised clustering is conducted on the extracted feature space to identify the brain (external and covert) states that respond to different cognitive demands. To understand the human state change, a state transition diagram is introduced, allowing visualization of connectivity patterns between every pair of states. We compute the transition probability between every pair of states to represent the relationships between the states. This state transition diagram is named as the Fuzzy Covert State Transition Diagram (FCOSTD), which helps the understanding of human states and human performance. We then apply FCOSTD on distracted driving experiments. FCOSTD successfully discovers the external and covert states, faithfully reveals the transition of the brain between states, and the route of the state change when humans are distracted during a driving task. The experimental results demonstrate that different subjects have similar states and inter-state transition behaviour (establishing the consistency of the system) but different ways to allocate brain resources as different actions are being taken.**

***Index Terms*— Human factors, state estimation, brain-computer interfaces (BCI), fuzzy neural networks.**

I. Introduction

BRAIN states can be defined as the intermediate representation of brain dynamics, which are generally recognized by electroencephalography (EEG), functional magnetic resonance imaging (MRI) or magnetoencephalography. Over the past decades, many brain-computer interface (BCI) systems based on EEG measurement technologies have been proposed for real-life applications [1]–[5]. Particularly, some passive BCI systems [4]–[11] have been developed for lapse detection, fatigue monitoring, attention evaluation, or accident prevention based on the characterization of the relationship between brain dynamics and behaviours, such as a subject's reaction time (RT) in response to a particular task or external stimuli. In previous studies, brain signals have been found to exhibit a highly stochastic temporal evolution associated with stable patterns of brain activity in response to external stimuli (external brain state) [1], [5], [9], [10], [12], [13]. However, the human brain also inherently exhibits latent mental processes (covert brain states). The underlying brain sources and their interactions might differ dramatically for different states [14]. These uncertainties in the human brain dramatically influence BCI performance [15]. Investigating the brain states, therefore, became a critical problem associated with BCI development and its performance.

Several approaches have been developed to resolve this problem. Study [16] proposed a decoding-based approach based on functional magnetic resonance imaging (fMRI) data to categorise unconscious and conscious mental states. In [17], the initial state of brain activity was identified by applying PCA on magnetoencephalography (MEG) data. This initial state then was used to trace brain activity in the state space as the subject received visual stimulus. These two studies considered spatial features that might not be able to discover time-dependent transition state. Taghia *et al.* [15] developed a system based on first-order Markov model to discover covert brain state by analysing functional connectivity from fMRI data. Kringelbach *et al.* [18] developed a framework that models the state transition in the brain based on the information extracted from neuroimaging data. This framework also used Markov model to observe state changes. Both [15] and [18] provide time-dependent approaches that enable the transition state being observed. The approaches for brain state discovery discussed so far rely on fMRI or neuroimaging data that offer only second-level observations on brain activity.

However, the alteration in mental state occurs in millisecond-level. And EEG device can record brain activity many times in a second, which provide more sophisticate information for brain state observation. Additionally, some covert states might be not directly measurable on its own

Manuscript received February 20, 2021; revised July 3, 2021 and September 29, 2021; accepted October 24, 2021. Date of publication November 8, 2021; date of current version December 8, 2021. This work was supported in part by the Australian Research Council (ARC) under Grant DP180100670, Grant DP180100656, and Grant DP210101093; in part by the Australia Defence Innovation Hub under Contract P18-650825; in part by the U.S. Office of Naval Research Global under Grant ONRG-NICOP-N62909-19-1-2058; in part by the U.S. Air Force Office of Scientific Research–Defence Science and Technology (AFOSR–DST) Australian Autonomy Initiative under Grant ID10134; and in part by the New South Wales (NSW) Defence Innovation Network and NSW State Government of Australia under Grant DINPP2019 S1-03/09. *(Corresponding author: Yu-Cheng Chang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of Taipei Veterans General Hospital under VGHIRB No. 2013-01-029BC.

Yu-Cheng Chang, Yu-Kai Wang, and Chin-Teng Lin are with the Faculty of Engineering and Information Technology, University of Technology at Sydney, Sydney, NSW 2007, Australia (e-mail: yu-cheng.chang@uts.edu.au).

Nikhil R. Pal is with the Electronics and Communication Sciences Unit (ECSU), Indian Statistical Institute (ISI), Kolkata 700108, India.

Digital Object Identifier 10.1109/TNSRE.2021.3126264

and should be identified through representative features of brain activity response to a task [19], [20]. For example, in case of a learning task, the brain resource is reallocated across the whole brain to response different cognitive demands [19], [21], [22]. Some states and their transitions occurring in brain resource reallocation may not be directly detectable by traditional BCI systems. Many existing methods [15]–[18] rely on neuroimaging to define the states and analyse their transitions. Therefore, exploring the covert state and its continuous transition from brain signals can benefit the development of BCI.

To obtain representative features of brain activities for covert state identification, we consider a fuzzy system (FS) to extract features from the brain dynamics. We assumed that every single fuzzy rule represents one specific brain pattern associated with a specific behaviour or latent mental process during the task. More specifically, the firing strength of a fuzzy rule can be the membership (degree) of a particular brain activity, which provides representative information to explicate the brain states [23]–[25]. In this study, the firing strength of fuzzy rules is used to provide intermediate information to explicate the brain states, and then an unsupervised density-based clustering method is exploited to define external states as well as the covert states of the brain. Another consideration of using the FSs is to defeat the uncertainty of brain activities. The fuzzification operation and if-then-rule architecture in FSs have been demonstrated to address uncertainty [23], [26]–[28]. The fuzzification operation translates the data into membership degrees using the Gaussian membership function and provides a tolerance for uncertainty, such as noise and variations in data [23], [26]–[29].

After obtaining representative information of brain activity, we use unsupervised learning to automatically identify covert and external brain states without external supervision. Thus, we used an unsupervised density-based clustering method to automatically identify brain states. One of the advantages of density-based clustering algorithms is that they do not require prior knowledge to determine the number of clusters, unlike centre-based clustering algorithms, such as K-means and K-medoids [30]–[32]. In addition, Density-based clustering algorithms calculate the local density of data so that arbitrarily shaped clusters can be detected, whereas most centre-based clustering algorithms might fail to identify clusters with an arbitrary shape [30], [32], [33].

To observe the changes between the brain states, we introduced the concept of Markov model [15], [34]–[37] to the captured brain states to obtain probability of transition between every pair of states. The transition probability of Markov model is not deterministic; it inherently cannot predict the future states from the present with certainty. However, this transition probability allows the user to find the possible feature states [15], [34]. Therefore, we can assign transition probability to brain states to generate moment-by-moment connectivity patterns, called the fuzzy covert state transition diagram (FCOSTD), which helps us to understand brain dynamics in relation to a cognitive task.

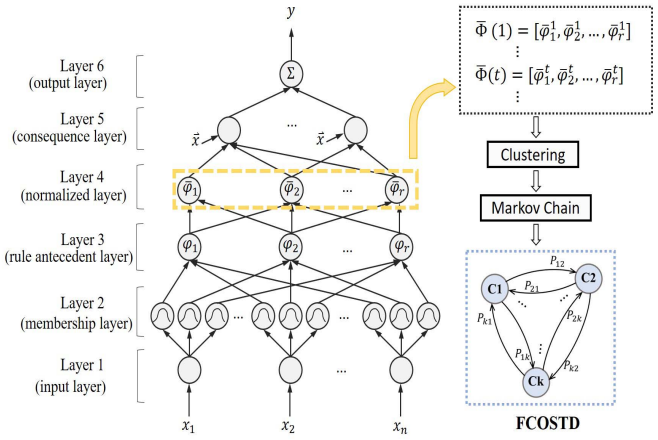


Fig. 1. The proposed FCOSTD. The fuzzy state identification includes a fuzzy neural network, clustering, and Markov chain, and then the FCOSTD can be generated.

The contribution of this research is threefold:

- A novel application of FSs that can encode brain activities as representative information by firing strength of fuzzy rules.
- The covert states of brain activity are identified by using an unsupervised clustering algorithm and linked to specific behaviour or latent mental process during the task.
- The brain (external and covert) states and transitions over time are visualised by the proposed FCOSTD.

The rest of the paper is organized in the following manner. Section 2 introduces the proposed framework, including the FNN, density-based clustering approach and FCOSTD. In Section 3, the FCOSTD is applied to EEG data collected from a series of distracted driving experiments of 11 subjects. Section 4 discusses and demonstrates the results of the three experiments. Finally, Section 5 describes the conclusions and future work concerning the application of the FCOSTD.

II. Methodology

Figure 1 interprets the generation of the FCOSTD based on EEG signals. Here the FNN plays an important role, providing plentiful information to carry out unsupervised clustering, brain state representation and state change visualization. The fuzzy inference system (FIS) of the FNN is used to encode brain activity based on frequency-domain EEG data.

A. Fuzzy Neural Network

To develop a brain-state-drift classifier that can overcome the subjective uncertainty of EEG signals and train the classifier with fewer segmental patterns, we exploited a SONFIN model [24], [28], a famous online-learning FNN, to classify incoming brain activity patterns. The SONFIN model includes an on-line structure learning algorithm to determine the number of fuzzy rules. In the initial stage of training, the SONFIN model has no fuzzy rules. Fuzzy rules are generated when the incoming training data cannot activate any existing rules.

This on-line structure learning algorithm ensures that each input datum is properly covered by a rule in the input space. Assume the input $X(t) = [x_1(t), \ldots, x_n(t)]^T \in \mathbb{R}^{n\times 1}$, and the SONFIN has learnt the structure with r rules. The FNN used in this study can be depicted as shown in Figure 1. The layers of the network, from the first to the fifth, realize a fuzzy-rule-based inference model of the following form:

$$\begin{aligned} &\text{Rule}: \text{ IF } x_1 \text{ is } A_1^i \text{ and } \ldots x_n \text{ is } A_n^i \\ &\text{THEN } w_i = a_0^i + a_1^i x_1 + \cdots + a_n^i x_n, \end{aligned} \tag{1}$$

where A_j^i is the i-th fuzzy set defined on the *j-th* input variable and $a_0^i, \ldots, a_n^i$ are the coefficients of the linear consequent of the rule. Note that the number of the consequent parameters corresponds one plus the dimension of the input variable X. To illustrate the structure of the FNN in detail, the FNN is introduced in this section layer by layer.

1) Layer 1 (Input Layer): The number of nodes in this layer equals the length of one input pattern vector. Each node, which performs no computation, transmits input data to the next layer. The outputs of the j-th node in this layer can be defined as $O_j^{(1)}) = x_j(t)$.

2) Layer 2 (Membership Layer): In layer 2, each node represents a fuzzy set and evaluates the membership value of an input variable from layer 1 with the corresponding membership function. We use a Gaussian membership function in this study. Thus, the output of each layer-2 node can be calculated as $O_{ij}^{(2)}) = \mu_j^i = exp[-(O_j^{(1)}) - m_j^i)^2/(\sigma_j^i)^2]$, where μ_j^i is the membership value corresponding to the *j-th* input variable of the i-th rule and m_j^i and σ_j^i are, respectively, the centre and the width of the Gaussian membership function.

3) Layer 3 (Rule Antecedent Layer): A node in layer 3 represents one fuzzy rule and receives corresponding membership values from the nodes in layer 2 to compute the firing strength value of the rule by performing a fuzzy AND operation. Here, fuzzy AND operation is utilized to calculate the node outputs, taking the algebraic product of the membership values of the atomic clauses associated with a rule. Given r rules, the output of each layer-3 node can be computed as $O_i^{(3)} = \varphi_i = \prod_j^n O_{ij}^{(2)}$.

4) Layer 4 (Normalization Layer): The size of this layer equals the size of layer 3. Each layer-4 node represents the normalized firing strength of the associated rule antecedent whose output is calculated by $O_i^{(4)} = \bar{\varphi}_i = O_i^{(3)} / \sum_k^r O_k^{(3)}$.

5) Layer 5 (Consequence Layer): Layer-5 nodes are called consequent nodes. Each node receives the output delivered from layer 4 and the input variables of layer 1. The output of layer 5 is obtained as $O_i^{(5)} = w_i \cdot O_i^{(4)} = \left(\sum_j^n a_j^i x_j + a_0^i\right) \cdot O_i^{(4)}$.

6) Layer 6 (Output Layer): A node of layer 6 corresponds to one output variable. Here, the nodes operate integration as a defuzzification process with $O^{(6)}) = y(t) = \sum_i^r O_i^{(5)}$.

The parameter-learning phase carries out supervised learning to optimize all the free FNN parameters. The parameters of fuzzy rules are learned via the gradient descent (GD) algorithm, and the structure is learned by an online clustering algorithm [28]. Considering the single-output case for simplicity, the error function E to be minimized is defined by $E = \sum_{t=1}^N \frac{1}{2}[y(t) - y_d(t)]^2$, where N is the number of training samples and y_d is the desired output for the layer-6 nodes.

B. Covert State Transition Diagram

After the optimised FNN is obtained, each rule antecedent component can learn a specific brain activity pattern corresponding to human behaviour. The outputs of each layer-4 node (normalised layer) represent the degree of matching levels between the input pattern and a particular brain activity pattern that has been stored as centres in a rule. Figure 1 shows the generation of the FCOSTD. Suppose that the FNN learned r rules from the frequency-domain EEG data, every observation $X(t)$ can be decomposed into r rules (stored patterns) and corresponding firing values (matching levels) that are encoded as a firing strength vector $\bar{\Phi}(t) = \left[\bar{\varphi}_1^t, \bar{\varphi}_2^t \ldots, \bar{\varphi}_r^t\right]$. We then utilise unsupervised clustering to categorise similar firing strength vectors into K clusters. According to the dominant label in each cluster, the K cluster can be defined as K states and connected to corresponding real behaviours or physical meanings by external labels of each cluster. The states here are covert states because they are identified based on the intermediate relationship of stored brain activity patterns. The relationships among the K covert states are then established to form a covert state transition diagram.

1) Density-Based Clustering Algorithm: We use the density peaks clustering algorithm (DPCA) [30] to cluster normalised firing strength vectors $\bar{\Phi}$s. Here, we define each strength vector $\bar{\Phi}$ as a data point. To find cluster centres for the whole set of $\bar{\Phi}$, DPCA computes the local density ρ_i and the distance δ_i to higher-density points for each datum point i. We use the Euclidean metric to calculate distances $d_{ij} = \left\|\bar{\Phi}(i) - \bar{\Phi}(j)\right\|$. The local density ρ_i is estimated by a Gaussian kernel function defined as

$$\rho_i = \sum_{j\neq i} exp\left[-\left(d_{ij}/d_c\right)^2\right], \tag{2}$$

where d_{ij} is the Euclidean distance between data points i and j, and d_c is the cut-off distance predefined by the user. Here, d_c is set according to Rodriguez's study [30], and a chosen d_c can satisfy that the average number of neighbours is approximately 20% of the total number of data points, e.g., $d_c = 0.07$. Given $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_N$, where N is the number of data points, and the δ_i can be determined by

$$\delta_i = \max_{j:\rho_j > \rho_i} d_{ij}. \tag{3}$$

Here δ_i represents the maximum distance between datum point i and other higher-density points. With these two quantities, we can define a γ_i value as follows:

$$\gamma_i = \rho_i \cdot \delta_i. \tag{4}$$

A high γ_i value of a datum point reflects that it simultaneously has crowding neighbours and is located far from other higher-density points. We can use this γ_i value to determine cluster centres. For other non-central points, clustering assignment is implemented by the nearest-neighbour approach.

Suppose Z be the set of points which have already been assigned clusters and we want x_i to assign its cluster. Let d_{ij} be the distance between x_i and $x_j \in Z$. Then the index of the nearest neighbour x_k of x_i is found as

$$k = \underset{j}{\operatorname{argmin}} \left\{ d_{ij} \,|\rho_j > \rho_i; x_j \in \mathrm{Z} \right\}. \tag{5}$$

Then x_i is assigned the same cluster label as that of x_k. Afterwards, each non-central point is assigned to the same cluster as its neighbour datum point of higher density.

2) Visualization of State Changes: To take a solid consideration for state representation of the proposed system, we must not only determine the covert states but also the relationships among them. We build a cover-state transition diagram taking transition probabilities into account to discover the temporal changes of the observed system via a Markov chain [34]–[36]. All connections among observable states are updated based on the state transition matrix. Given K states found by the DPCA, the state transition matrix can be defined as

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1K} \\ \vdots & \ddots & \vdots \\ p_{K1} & \cdots & p_{KK} \end{bmatrix}. \tag{6}$$

The columns of the matrix refer to the state for the next time point, and the rows of the matrix refer to the current state. Each transition probability p_{ij} between two states is estimated by the maximum likelihood approach, which is $p_{ij} = T_{ij}/L$, where T_{ij} is the number of direct transitions from state i to state j and L is the total number of times the system arrived at the state i, irrespective of how it has arrived. Thus $\sum_{j=1}^{K} p_{ij} = 1; i = 1, 2, \ldots, K$. Once the state transition matrix is obtained, we can observe the path of transition from state to state as the observed system reacts to the stimuli or tasks.

III. Experiment

A. Experimental Design

The human brain is a complex system, and brain dynamics change rapidly depending on stimuli, tasks or emotion. To demonstrate the feasibility of the proposed FCOSTD, the EEG signals recorded from the distracted driving experiment were selected because the participant must switch their attention between the designed tasks over time. We used a 10-20 electrode placement cap with a 32 Ag/AgCl electrodes (NeuroScan, NeuroScan Inc., Herndon, VA, USA) to collect EEG signals at 500-Hz sampling rate. All electrodes of the EEG cap were referenced to linked two mastoids of the participant, and a single ground electrode was attached to the forehead. All references and ground are built in the device. The contact points between the EEG electrodes and the scalp were applied with conductive gel to reduce the contact impedance to less than 5 kΩ.

The experiment was designed in a driving simulator on a dynamic 6-degree-of-freedom motion platform with 360° driving scenes rendered on the seven LCD projectors [9], [38]. The simulator provides authentic visual and kinesthetic stimuli of the driving situation on a highway to the participants. The participants experience real-time kinematic feedback from the motion platform regarding their operations, such as turning the steering wheel. The simulator emulates car cruising at a fixed speed of 100 km/hr on a highway scene in the night throughout the entire experiment. There are two tasks: one driving task and a mental calculation task. In the driving task, the car deviates to either the left or right from the cruising lane randomly, and the participants are asked to turn the steering wheel to control the car back to the cruising lane immediately. In the mental calculation task, the participants need to verify a two-digit addition equation (i.e., whether the equation is correct or incorrect) displayed on the front screen when the task comes in. The numbers of correct and incorrect equations are the same, and the difficulty level is set the same as well. There are two buttons mounted on the left and right sides of the steering wheel. The participants can easily use the right thumb to press the button mounted on the right side of the steering wheel to report the correctness of the mathematical equation. To evaluate behaviour performance, the latency between even on-set to proper response is defined as reaction time (RT). For the driving task, the RT is the latency between the deviation on-set and turning the steering wheel. Similarly, the latency between the presentation of the equation and the button press is the RT of the mathematical task.

Thirteen neurologically healthy volunteers aged 20–28 years participated in a distracted driving experiment. Continuous EEG data were recorded during the time a participant was involved in the experiment. All participants were required to have a driver's license and good driving habits. This experiment was conducted at National Chiao Tung University, Hsinchu, Taiwan [9], [10]. Every participant had two training sessions before the experiment to acquaint themselves with the two tasks and the virtual-reality environment. After two training sessions, every participant had four 15-min experimental sessions separated by 10-min rests. Each participant was recorded as many as 100 trials in each session to ensure that the number of tasks was adequate for statistical analysis. A more detailed explanation of the designed distracted driving can be found in our previous studies [9], [38].

B. EEG Data Preprocessing

In the preprocessing phase, the continuous EEG recordings are filtered using a zero-phase finite impulse response bandpass filter, which has been widely used for EEG signal prepossessing [39]–[41]. We set the lower cut-off frequency to 1 Hz, and the higher cut-off frequency is 50 Hz. The selected cut-off frequencies enable the filter to capture all the variations in event-related spectral perturbation of collected EEG signals in the designated bandwidth and filter out the noise from the power line, movement-related artifacts, and sweat artifacts. Afterwards, the continuous EEG data were segmented into a set of epochs, of which the length was from -1 second (prior to the first stimuli on-set) to 5 seconds (following the first stimuli on-set) to comprise all relevant activations. We manually eliminated those epochs with large eye movements, body movements, and amplifier saturation. Other wide variety of artifacts are then removed via independent component analysis

TABLE I

THE PERFORMANCE OF FNN AND CLUSTERING ON THE DISTRACTED DRIVING EXPERIMENTS. THE NUMBER IN THE FIRST ROW REPRESENTS THE INDEX OF THE SUBJECTS. THE FIRST AND SECOND ROWS PRESENT THE NUMBER OF RULES (r) THAT FNN GENERATED AND ITS RECOGNITION RATE (RR), RESPECTIVELY. THE PERFORMANCE OF CLUSTERING IS EVALUATED BY USING THE SILHOUETTE SCORE (**SI**) AND FOWLKES-MALLOWS SCORE (**FI**), AS SHOWN IN THE THIRD AND FOURTH ROW, RESPECTIVELY

	subject_1	subject_2	subject_3	subject_4	subject_5	subject_6	subject_7	subject_8	subject_9	subject_10	subject_11	subject_12	subject_13
r	14	14	15	12	17	13	16	13	15	19	15	12	15
RR (%)	93.36	93.87	93.74	87.82	96.09	89.99	92.73	94.58	93.42	93.55	93.40	94.47	95.08
SI	0.56	0.53	0.59	0.52	0.39	0.61	0.43	0.39	0.5	0.36	0.47	0.57	0.61
FI	0.56	0.49	0.58	0.57	0.67	0.56	0.53	0.62	0.54	0.53	0.55	0.52	0.57

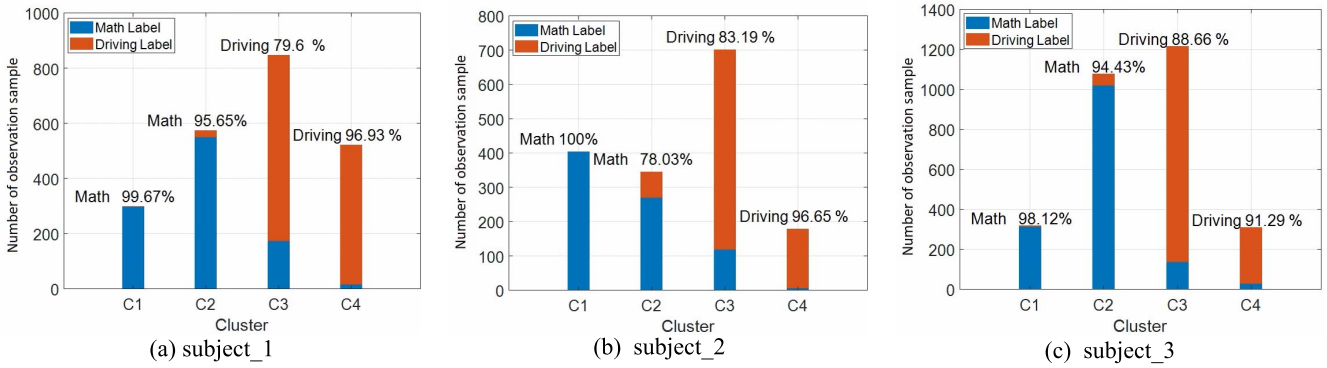
(ICA) [42], [43]. The independent component analysis (ICA) models the EEG signals from different brain sources into linear aggregation. The electrical activities of all individual EEG channels and artifactual sources can be effectively isolated by the ICA [42], [43]. Therefore, the artifacts, such as eye blinks and muscle activities, can be detected and removed easily. Once the artifact-free EEG data are produced, it was segmented into a 1.2-second interval from the time a task is on-set, and we call it a trial. Then, every trial is further divided into 17 observation samples with overlapping 400-ms windows advancing in 50-ms steps. Every observation sample was transferred from the time domain to the frequency domain using the fast Fourier transform (FFT). We normalised the EEG power spectra of each frequency by subtracting from the mean of each spectrum in the baseline. The baseline was taken from the pre-stimulus period (from -1 second to 0 second). We selected six components, the frontal, central, parietal, occipital, left motor and right motor regions, for further processing. The mean powers in the delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), and low beta (13–20 Hz) bands were then extracted from the selected components. Note that previous studies [1], [9], [23], [38], [44] have demonstrated that the brain dynamics in these selected components are highly associated with visual perception, visual processing, motor control, attention and planning. The findings in [10] and [47] also suggest that human mental states can change significantly within 1.2 seconds after the task is on-set during the driving. For the training process of the FNN, all 17 observation samples from one trial are assigned one single label, either driving or math. For a single participant, the size of input data fed to FNN equals to Number of Trails $\times$ 17 observation samples. For every single observation sample, the dimension of feature is 24 (4 bands $\times$ 6 components).

C. Experimental Results

The FNN, in this study, is trained subject-by-subject because different subjects have various brain activity patterns. The output of FNN is binary. 0 represents driving and 1 is math. We use mean squared error (MSE) as the loss function for training. The learning rate for update rule of gradient descent algorithm is set to 0.001, and the learning epoch is set to 100. The single-subject performance of all experiments is shown in Table I. The second and third rows, respectively, present the number of rules (r) that FNN generated and its recognition rate. The fourth and fifth rows present the performance of clustering on firing strength vectors $\bar{\Phi}$ extracted from the trained FNN. The performance of clustering is evaluated by using the Silhouette score (**SI**) [51] and the Fowlkes-Mallows score (**FI**) [46]. The Silhouette score in Table I is the mean value across the estimated compactness values of each cluster, while the Fowlkes-Mallows score shown in the next row (in its right lower box) evaluates the geometric mean of the precision and the recall. The number of clusters for all subjects is determine by γ_i value in formula (4), more details are described in section IV-*A*.

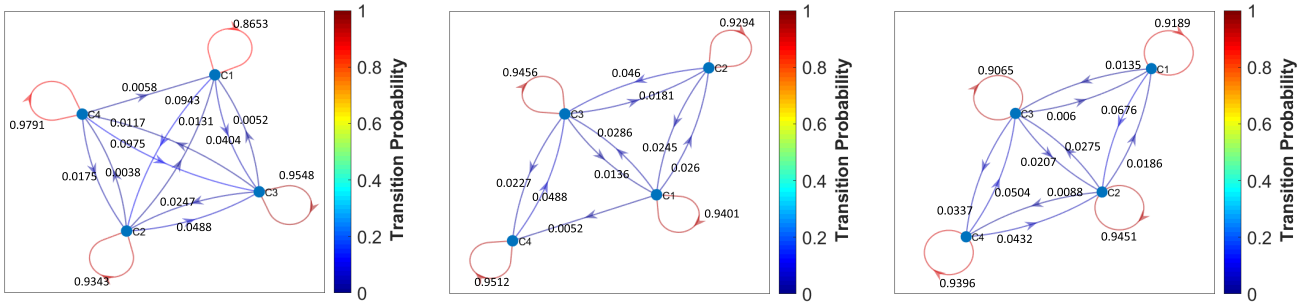
The FNN trained with data from subject_5 has the highest recognition rate of 96.09%, but the silhouette score of its clustering result stands at a relatively low level, with a value of 0.3949. On the other hand, the recognition rate for the experiment with subject_6 is 89.99%, which is in the second-lowest place, but the Silhouette score is 0.6096, which is in the first place. These two subjects have low performance either in terms of recognition rate or in terms of silhouette score. Subject_4 has the lowest recognition rate at 87.82%, but its Silhouette and Fowlkes-Mallows scores are 0.52 and 0.57 that are both acceptable clustering performance. The performances of both subject_7 and subject_8 are similar to that of subject_5; its recognition rate is relatively high, but the silhouette score is relatively low. Subject_9, subject_10 and subject_11 all perform well in terms of recognition rate as well as Fowlkes-Mallows Score, but the Silhouette scores of subject_10 and subject_11 both are lower than 0.5.

We categorised the driving behaviour as three types: distracted driver (poor driving performance), dynamic driver (medium performance in driving and mental calculation) and focused driver (best driving performance). We choose the experiments of other subjects whose recognition rate and clustering performance both are relatively higher, e.g., the recognition rate is over 90% and the silhouette scores is over 0.5%, as representative subjects for the analysis of the three driving behaviour types. Subject_1 is a distracted driver, has shortest RT (1.596 sec) in doing math, but longest RT (0.775 sec) in driving, which is a distracted driving behaviour, as shown in Table II. Subject_2 is a focused driver, has a relatively low RT (0.663 sec) in driving, but its RT of doing math is the longest value (1.709 sec), which is a case of highly focusing on driving. Subject_3 is a dynamic driver, performed well both in driving and doing math, the RTs is 0.648 sec and 1.53 sec, respectively. We emphasize that the analysis



(a) subject_1 (b) subject_2 (c) subject_3

Fig. 2. (a), (b) and (c) present clustering results of subject_1, subject_2 and subject_3, respectively. The math label and driving label refer to the mental calculation task and driving task, respectively. The percentage over each bar represents the sharing rate of the dominating label.



(a) subject_1 (b) subject_2 (c) subject_3

Fig. 3. (a), (b) and (c) present state transition diagram of subject_1, subject_2 and subject_3, respectively. The digits adjacent to each connection line are the transition probabilities.

TABLE II
THE RT OF SUBJECT_1, SUBJECT_2 AND SUBJECT_3 DURING THE MENTAL CALCULATION TASK AND DRIVING TASK. UNIT: SECONDS

Task	Driving	Mental calculation
Subject_1	0.775	1.596
Subject_2	0.663	1.709
Subject_3	0.648	1.530

that follows would be equally valid for other subjects, but we choose subject_1, subject_2, and subject_3 so that the outcome of this study becomes easier to visualize by any reader.

D. Distracted Driver (Poor Driving Performance)

Figure 2 shows the content of each clustering, and each index on the horizontal axis represents a cluster, i.e., Cn, n = 1, 2, 3, and 4, represent the n-*th* cluster. There are four clusters generated by the collected data from the subject_1, as shown in Figure 2 (a). Each cluster has its own dominating label, for which the occupancy rate is the greatest. The math label (the blue part of a bar) indicates that the participant doing or paying attention to the mental calculation task, while the driving label (the saffron part of a bar) represents the driving task. Table III reveals the distribution percentage of each label in different clusters.

In Figure 2 (a), the clusters C1 and C2 are dominated by math labels, which means that for most of the trails included in these two clusters this subject allocates most of the brain resources to math tasks. Particularly, there are more samples labelled as solving math in C2. Based on the label distribution, we define C1 and C2 as a pure calculating and a covert state from mental calculating to driving, respectively. There are more samples labelled as driving in C3 and C4 (ref. Figure 2 (a)). Here, the state (cluster) C4 is almost a pure state representing driving behaviour due to the 96.93% occupancy rate of the driving label. The state C3 consists of 20.4% math labels and 79.60% driving labels. Thus, it is reasonable to consider C3 a covert state of driving but relatively close to the calculating behaviour.

Figure 3 (a) is a state transition diagram for subject_1. The four solid circles on the diagram correspond to the four different clusters/states, as shown in Figure 2 (a). The directed connection between every two states represents the potential transition with the probability indicated therein. Table IV summarises the transition probabilities between every two states. For example, between C1 and C2, there are two connection lines: one starts at C1 and ends at C2, and the other is the opposite direction. C1 and C2 exhibit bidirectional transitions. More specifically, the conditional probability of transition from C1 to itself (C1) is 0.8653 and that from C1 to C2 and C3 are 0.0943 and 0.0404, respectively. This probability of transition from C1 to any other state including itself does not depend on the state of the system before it came to C1 (i.e., it is a Markov chain). It is worth noting that there is no transition from C1 to C4 which is very reasonable because

TABLE III
THE DISTRIBUTION PERCENTAGE OF EACH LABEL IN ALL CLUSTERS OF SUBJECT_1, SUBJECT_2 AND SUBJECT_3

UNIT: %	Subject_1		Subject_2		Subject_3	
Cluster	Math Label	Math Label	Driving Label	Driving Label	Driving Label	Driving Label
C1	99.67	0.33	100	0	98.12	1.88
C2	95.65	4.35	78.03	21.97	94.43	5.57
C3	20.4	79.6	16.81	83.19	11.34	88.66
C4	3.07	96.93	3.35	96.65	8.71	91.29

TABLE IV
TRANSITION PROBABILITY BETWEEN EACH TWO IDENTIFIED STATES OF SUBJECT_1, SUBJECT_2 AND SUBJECT_3

	Subject_1				Subject_2				Subject_3			
States	C1	C1	C4	C2	C3	C4	C2	C3	C4	C2	C3	C4
C1	0.8653	0.0943	0.0404	0	0.9401	0.026	0.0286	0.0052	0.9189	0.0676	0.0135	0
C2	0.0131	0.9343	0.0488	0.0038	0.0245	0.9294	0.046	0	0.0186	0.9451	0.0275	0.0088
C3	0.0052	0.0247	0.9584	0.0117	0.0136	0.0181	0.9456	0.0227	0.006	0.0207	0.9396	0.0337
C4	0.0058	0.0175	0.0975	0.8791	0	0	0.0488	0.9512	0	0.0432	0.0504	0.9065

C1 is a pure Math-state and C4 is almost a pure driving state.

For Subject 1, although there is a very low probability of transition from C4 to C1, there is no direct transition C1 to C4. A natural question comes – are there indirect transitions from C1 to C4. Further analysis of the data shows that there was no indirect transition from C1 to C4 via C2 or C3. Here, by indirect transition we refer to the situation when the system switches from C1 to C2 and then C2 to C1 in two successive time instants, i.e., without making any transition from C2 to itself. This again emphasizes the stability of the identified brain states. It is interesting to note that C2 is not a pure state, but this is a very stable internal (Covert) state because the conditional probability of transition from C2 to itself is 0.9343. Consequently, the conditional probability of transitions to C1, C3 and C4 are very small, 0.0131, 0.0488, and 0.0038, respectively. Similarly, the covert state C4 is a very stable one, for which the conditional probability of transition from C4 to itself is the highest, 0.9791.

E. Focused Driver (Best Driving Performance)

The number of clusters used for subject_2, as shown in Figure 2 (b), is also four, which is the same as that of subject_1. Table III shows the distribution percentage of each label in all four clusters. The samples categorised in C1 are all math labels. C2 has math and driving labels mixed at 78.03% and 21.97%, respectively. In C3, the driving label occupies 86.91%, while the math label occupies 16.81%. C4 contains almost pure driving labels of approximately 96.55%. Therefore, we define C1 as a pure calculating behaviour, C2 as the covert state from calculating to driving behaviour, C3 as from driving to calculating to driving behaviour, and C4 representing pure driving behaviour. Figure 3 (b) is the corresponding state transition diagram for subject_2.

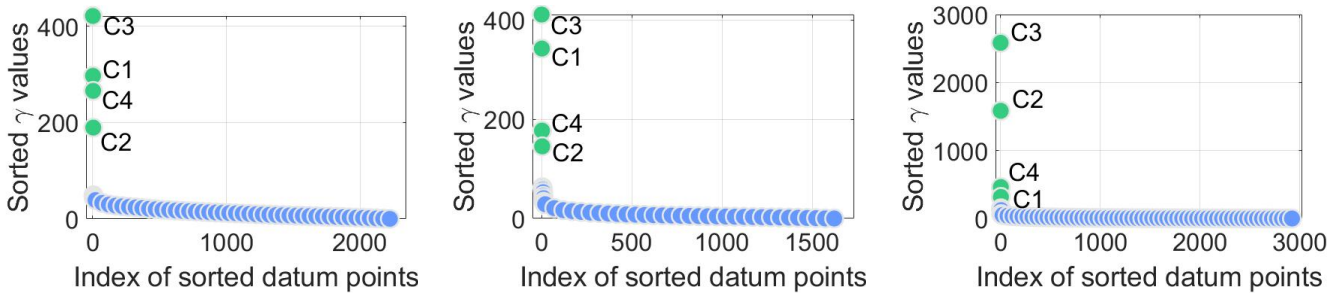
Table IV shows the conditional transition probabilities between every two states for subject_2. Like Subject_1, the high conditional probability of transition from any of the four states to itself (more than 0.92) suggests that each of them represents a valid state. Unlike Sublect_1, for Subject_2, there is no transition from C4 to C1 and a very weak transition probability from C1 to C4. This is just the opposite of Subject_1. A possible reason for this may the fact that for Subject_2, C4 is almost a pure driving state and C1 is almost a pure math sate, but C4 is more strongly represented (with much greater number of trials mapped in C4) than C1. On the other hand, for Subject_1, C1, a pure math state, is much more strongly represented than the almost pure driving state C4. So, from these two subjects, what we observe is that between two pure states, if one is represented by a smaller number of trials, then there is a low chance of transition from the weaker one to the stronger one. Like Subject 1, there is no indirect transition between C4 and C1 (ref. Figure 3 (b)).

IV. DISCUSSION

A. Determining the Number of Clusters

In Figure 4 the panels (a), (b) and (c) depict the γ values of subject_1, subject_2, and subject_3 in decreasing order, respectively. Each solid circle represents one data point, in which the label indicates the associated cluster. Figure 4 lists only samples with the top fifty percentages of sorted γ values. The γ value can be used as a criterion to select cluster centres. The sorted γ values approximately follow a power low, as shown in Figure 4. There are four isolated points strikingly distinctive from those who are close to the horizontal axis. As explained earlier, if a point has a high γ_i value (Ref. formula (4)), then that point has high-density crowding neighbours and is located significantly away from other higher-density points. Therefore, these distinctive points are selected as centres for the clustering.

For example, in Figure 4 (a), the first point on the top has the highest value of the product of highest density and longest distance from other relatively high-density points. Thus, this sample is defined as a centre. The rank-two sample is defined as a second centre due to its second highest γ implying that



(a) Subject_1 (b) Subject_2 (c) Subject_3

Fig. 4. Distribution of sorted γ values in descending order. Only those samples with the top 50 percent high γ values are shown in (a), (b) and (c). The indices on each dot indicate to which cluster belongs.

it is a high-density point as well as distant from other high-density points. Followed by the third and fourth centres, they both are high γ value samples. The remaining non-isolated samples in Figure 4 (a), cannot be the centres because the γ values of these instances are relatively smaller than those in the selected four isolated cases. Then we take these four data points as cluster centres and form four clusters around them for subject_1, as shown in Figure 2. These four top-ranked samples belonging to C3, C1, C4 and C2 correspond to clusters 3, 1, 4 and 2, respectively, as labelled in Figure 2 (a). Analysing the clusters, we have exhibited that these four clusters correspond to four different types of brain states associated with handling driving tasks and mathematical calculation tasks. The same strategy is also applied to subject_2 and subject_3 and surprisingly for each of them the system discovers four clusters representing four brain states.

B. State Change Analysis

We leveraged FCOSTD to explore the changes in human states through brain dynamics during driving. Note that the changes in human state can be linked with human behaviour as humans perform driving tasks or mental calculations during distracted driving. In this study, we demonstrated the changes in human states by brain dynamics and listed two different results in which two participants applied different strategies when facing a high cognitive task (mental calculation) during driving. The brain dynamics evolve continuously, and the samples are extracted only over 1200 ms. The driving task mainly involves motor reflection in which the visual perception and motor controlling resources are primarily involved [38], [47]. Meanwhile, visual perception, visual memory, and executive functions in the human brain are highly associated with the process of mental calculation [38], [48]. Based on the findings of cognitive functions in the previous studies [9], [10], [38], here we investigate the state changes over 1200 ms period and leverage the behaviour performance (RTs) to infer the state changes while performing math and driving.

Table V illustrates the switching time of each state. Let us elaborate on how the switching time is computed. Each trail is of 1200 ms (i.e., 1.2 sec). Each trial is divided into 17 segments, each segment is of length 400ms with an overlap of 350 ms between two successive segments. Suppose the k^{th} segment ($k < 17$) of a trial is in state Cx ($x = 1, 2, 3, 4$), and the $(k+1)^{th}$ segment of that trial is in some other state Cy ($y \neq x$) then the switching time for that trial for state Cx is $(1.2/17) * (k+1)$. In Table V we report the average value of the switching time for each state along with its standard deviation. Each state is counted across all trials, and the statistical data are shown in Table V.

When subject_1 performs mental calculation during driving, the subject's present state switches quickly from the pure mental calculation (C1) to covert states (C2 or C3); the average time of switching from C1 to other states is 0.4228 ms (Table V). Apart from the mental calculation, the low switching time from C4 to other states implies that subject_1 can quickly allocate brain resources to finish the driving task and reallocate the whole brain resource for the next task of mathematical calculation. The average switching time from C4 to other states is 0.4291 ms. Figure 3 (a) implies the same phenomena. The diagram shows a fully connected graph except for the connection from C1 to C4; thus, the brain states of subject_1 can switch from C1 to C2 or to C3 and from C4 to other states but from C1 to C4. Subject_1, therefore, needs more brain resources to switch the brain state to take action when the car deviates from the lane.

In contrast, subject_2 exhibits significantly different behaviour and states compared to those of subject_1. The states of subject_2 change less frequently, and the average RT is lower when performing mental calculation. Subject_2 allocates higher brain resources for mental calculation and has a delay in switching between states compared to subject_1. Its transition diagram also reflects this observation, as shown in Figure 3 (b). The transition probabilities from C4 to C1 or to C2 as well as that from C2 to C4 are all zero, which implies that the state cannot change directly from C4 to C1; it needs to move to C3 or C2 first and then from there to reach C1. More specifically, subject_2 needs more time to prepare to perform mental calculations. Thus, subject_2 exhibited a lower reaction time performance than subject_1. Due to the high resources used for mental calculations, it is relatively slow when the states change from C1 to C4 when subject_2 performs mathematical calculation. However, the average value of subject_2's RT is smaller than that of subject_1's RT. This is because turning the steering wheel is sensory decision-making;

TABLE V
SUMMARY OF SWITCHING TIME FROM ANY GIVEN STATE TO ALL OTHER STATES

Unit: sec	*Subject_1*		*Subject_2*		*Subject_3*	
Switching Time	**AVG**	**STD**	**AVG**	**STD**	**AVG**	**STD**
C1	0.4228	0.2562	0.6258	0.3458	0.6476	0.3597
C2	0.7168	0.3297	0.6048	0.3451	0.6188	0.3402
C3	0.7672	0.3346	0.6118	0.3395	0.5904	0.3379
C4	0.4291	0.2919	0.7581	0.3164	0.8572	0.296

it uses lesser brain resources and can take action in response to car deviation. This phenomenon can also be observed in the driving task of subject_3.

Subject_3 prepares to correct the deviation of the car faster than other subjects; hence, the distractive driving state (C3) appears at an average switch time of 0.5904 sec, which is earlier than that of subject_2 and subject_1. This earlier switching of the driving state results in faster average reaction time in the driving task. Apart from the driving task, subject_3 performed mental calculations with a 1.53-sec average RT (ref. Table II); the states remained in C2 for many trials. This is suggestive of the fact that subject_3 did not allocate much brain resources for mental calculations and has a good capability to address driving and mental calculations simultaneously. In Figure 3 (c), the transition diagram shows that the transition probabilities from C1 to C4 and C4 to C1 are both zero, which indicate that the two pure states, the mental calculation (C1) and driving task (C4), cannot directly switch from one to the other. However, the transition probability from C4 to C2 is 0.0432, higher than that of both subject_1 and subject_2. This phenomenon implies that subject_3 can quickly allocate brain resources for mental calculations from driving tasks. Accordingly, we may infer that the load on the brain of subject_3's is not as heavy as that of subject_2 and subject_1 during the whole experiment. Consequently, subject_3 can switch the brain states faster than others.

V. CONCLUSION

In the current study, we proposed FCOSTD which provides an effective mechanism for discovery and representation of the covert states and the transitions between states. We applied FCOSTD to a distracted driving experiment. The proposed system is capable of extracting useful features from the dynamic patterns of brain signals using an FNN. The FNN, leveraged by the fuzzy-inference mechanism, was successfully used for feature encoding from the EEG data. We adopted this approach because the inference mechanism of an FNN is capable of representing/modelling human brain activities associated with varying behaviours (RTs). These fuzzy-inference-based features are then clustered to identify external and covert states of human brain during the experiment involving distracted driving. The clustering results are used to define and interpret the states in terms of physical behaviours. Afterwards, the relationship between different states is expressed by the transition probability diagram. Because of the links between pairs of states along with the link weights representing the conditional transition probabilities, brain dynamics can be easily visualized. We found that covert states do exist in the brain when the subject is responding to on-set tasks during the experiment. Moreover, the FCOSTD also provides a mechanism for describing state changes with their corresponding probabilities and the Markov chain. Through the Markov chain, the state changes can be easily visualized, which allows us to understand how the brain resource was allocated as different types of actions were taken up by a subject. In the near future, we want to exploit the inherent interpretabliltiy of if-then-rule architecture of FNN to explain brain state changes in a different perspective, and the states discovered by the FCOSTD can be used as one major factor for improving human-machine autonomous systems.

REFERENCES

[1] C.-T. Lin *et al.*, "EEG-based brain-computer interfaces: A novel neurotechnology and computational intelligence method," *IEEE Syst., Man, Cybern. Mag.*, vol. 3, no. 4, pp. 16–26, Oct. 2017, doi: 10.1109/MSMC.2017.2702378.

[2] 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," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 400–406, Feb. 2018, doi: 10.1109/TNSRE.2018.2790359.

[3] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in brain-computer interfaces: A review," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1753–1762, Oct. 2017, doi: 10.1109/TNSRE.2016.2627016.

[4] C.-T. Lin, C.-H. Chuang, Y.-C. Hung, C.-N. Fang, D. Wu, and Y.-K. Wang, "A driving performance forecasting system based on brain dynamic state analysis using 4-D convolutional neural networks," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4959–4967, Oct. 2021, doi: 10.1109/TCYB.2020.3010805.

[5] C.-T. Lin *et al.*, "Exploring the brain responses to driving fatigue through simultaneous EEG and fNIRS measurements," *Int. J. Neural Syst.*, vol. 30, no. 1, Jan. 2020, Art. no. 1950018, doi: 10.1142/S0129065719500187.

[6] S.-Y. Dong, B.-K. Kim, and S.-Y. Lee, "EEG-based classification of implicit intention during self-relevant sentence reading," *IEEE Trans. Cybern.*, vol. 46, no. 11, pp. 2535–2542, Nov. 2016, doi: 10.1109/TCYB.2015.2479240.

[7] 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, Apr. 2014, doi: 10.1109/TBCAS.2014.2316224.

[8] S. P. Patel, B. P. Patel, M. Sharma, N. Shukla, and H. M. Patel, "Detection of drowsiness and fatigue level of driver," *Int. J. Innov. Res. Sci. Technol.*, vol. 1, no. 11, p. 6, 2015.

[9] Y. Wang, T.-P. Jung, and C.-T. Lin, "EEG-based attention tracking during distracted driving," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 6, pp. 1085–1094, Nov. 2015, doi: 10.1109/TNSRE.2015.2415520.

[10] Y.-K. Wang, S.-A. Chen, and C.-T. Lin, "An EEG-based brain–computer interface for dual task driving detection," *Neurocomputing*, vol. 129, pp. 85–93, Apr. 2014, doi: 10.1016/j.neucom.2012.10.041.

[11] K. Huang, C. Chuang, Y. Wang, C. Hsieh, J. King, and C. Lin, "The effects of different fatigue levels on brain–behavior relationships in driving," *Brain Behav.*, vol. 9, no. 12, Sep. 2019, Art. no. e01379, doi: 10.1002/brb3.1379.

[12] L. Lotte *et al.*, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, 2018, Art. no. 031005, doi: 10.1088/1741-2552/aab2f2.

[13] K. Katahira, Y. Yamazaki, C. Yamaoka, H. Ozaki, S. Nakagawa, and N. Nagata, "EEG correlates of the flow state: A combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task," *Frontiers Psychol.*, vol. 9, p. 300, Mar. 2018, doi: 10.3389/fpsyg.2018.00300.

[14] C.-T. Lin, M. Nascimben, J.-T. King, and Y.-K. Wang, "Task-related EEG and HRV entropy factors under different real-world fatigue scenarios," *Neurocomputing*, vol. 311, pp. 24–31, Oct. 2018, doi: 10.1016/j.neucom.2018.05.043.

[15] J. Taghia *et al.*, "Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition," *Nature Commun.*, vol. 9, no. 1, pp. 1–19, Jun. 2018, doi: 10.1038/s41467-018-04723-6.

[16] J.-D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," *Nature Rev. Neurosci.*, vol. 7, no. 7, pp. 523–534, Jul. 2006, doi: 10.1038/nrn1931.

[17] A. T. Baria, B. Maniscalco, and B. J. He, "Initial-state-dependent, robust, transient neural dynamics encode conscious visual perception," *PLOS Comput. Biol.*, vol. 13, no. 11, Nov. 2017, Art. no. e1005806, doi: 10.1371/journal.pcbi.1005806.

[18] M. L. Kringelbach *et al.*, "Dynamic coupling of whole-brain neuronal and neurotransmitter systems," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 17, pp. 9566–9576, Apr. 2020, doi: 10.1073/pnas.1921475117.

[19] A. Yousefi *et al.*, "Decoding hidden cognitive states from behavior and physiology using a Bayesian approach," *Neural Comput.*, vol. 31, no. 9, pp. 1751–1788, Sep. 2019, doi: 10.1162/neco_a_01196.

[20] M. L. Kringelbach and G. Deco, "Brain states and transitions: Insights from computational neuroscience," *Cell Rep.*, vol. 32, no. 10, Sep. 2020, Art. no. 108128, doi: 10.1016/j.celrep.2020.108128.

[21] T.-T.-N. Do, C.-H. Chuang, S.-J. Hsiao, C.-T. Lin, and Y.-K. Wang, "Neural comodulation of independent brain processes related to multitasking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1160–1169, Jun. 2019, doi: 10.1109/TNSRE.2019.2914242.

[22] T.-T.-N. Do, Y.-K. Wang, and C.-T. Lin, "Increase in brain effective connectivity in multitasking but not in a high-fatigue state," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 3, pp. 566–574, Sep. 2021, doi: 10.1109/TCDS.2020.2990898.

[23] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 9, pp. 2044–2055, Sep. 2012, doi: 10.1109/TCSI.2012.2185290.

[24] Y.-T. Liu, Y.-Y. Lin, S.-L. Wu, C.-H. Chuang, and C.-T. Lin, "Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 2, pp. 347–360, Feb. 2016, doi: 10.1109/TNNLS.2015.2496330.

[25] A. Jafarifarmand, M. A. Badamchizadeh, S. Khanmohammadi, M. A. Nazari, and B. M. Tazehkand, "A new self-regulated neuro-fuzzy framework for classification of EEG signals in motor imagery BCI," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1485–1497, Jun. 2018, doi: 10.1109/TFUZZ.2017.2728521.

[26] I. Caylak, E. Penner, and R. Mahnken, "A fuzzy uncertainty model for analytical and numerical homogenization of transversely fiber reinforced plastics," *PAMM*, vol. 19, no. 1, Nov. 2019, Art. no. e201900356, doi: 10.1002/pamm.201900356.

[27] I. Couso, C. Borgelt, E. Hullermeier, and R. Kruse, "Fuzzy sets in data analysis: From statistical foundations to machine learning," *IEEE Comput. Intell. Mag.*, vol. 14, no. 1, pp. 31–44, Feb. 2019, doi: 10.1109/MCI.2018.2881642.

[28] C.-F. Juang and C.-T. Lin, "An online self-constructing neural fuzzy inference network and its applications," *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 1, pp. 12–32, Feb. 1998, doi: 10.1109/91.660805.

[29] T. K. Reddy, V. Arora, L. Behera, Y.-K. Wang, and C.-T. Lin, "Multiclass fuzzy time-delay common spatio-spectral patterns with fuzzy information theoretic optimization for EEG-based regression problems in brain–computer interface (BCI)," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 10, pp. 1943–1951, Oct. 2019, doi: 10.1109/TFUZZ.2019.2892921.

[30] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014, doi: 10.1126/science.1242072.

[31] L. C. C. Heredia and A. R. Mor, "Density-based clustering methods for unsupervised separation of partial discharge sources," *Int. J. Electr. Power Energy Syst.*, vol. 107, pp. 224–230, May 2019, doi: 10.1016/j.ijepes.2018.11.015.

[32] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, 2010, doi: 10.1016/j.patrec.2009.09.011.

[33] K. Pal, N. R. Pal, J. M. Keller, and J. C. Bezdek, "Relational mountain (density) clustering method and web log analysis," *Int. J. Intell. Syst.*, vol. 20, no. 3, pp. 375–392, Mar. 2005, doi: 10.1002/int.20071.

[34] W. Jamal, S. Das, I.-A. Oprescu, and K. Maharatna, "Prediction of synchrostate transitions in EEG signals using Markov chain models," *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 149–152, Feb. 2015, doi: 10.1109/LSP.2014.2352251.

[35] P. Salvador, A. Nogueira, and R. Valadas, "Markovian models for medical signals on wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2009, pp. 1–5, doi: 10.1109/ICCW.2009.5208090.

[36] B. Kemp and H. A. C. Kamphuisen, "Simulation of human hypnograms using a Markov chain model," *Sleep*, vol. 9, no. 3, pp. 405–414, Sep. 1986, doi: 10.1093/sleep/9.3.405.

[37] S. Ryali *et al.*, "Temporal dynamics and developmental maturation of salience, default and central-executive network interactions revealed by variational Bayes hidden Markov modeling," *PLOS Comput. Biol.*, vol. 12, no. 12, Dec. 2016, Art. no. e1005138, doi: 10.1371/journal.pcbi.1005138.

[38] Y.-K. Wang, T.-P. Jung, and C.-T. Lin, "Theta and alpha oscillations in attentional interaction during distracted driving," *Frontiers Behav. Neurosci.*, vol. 12, p. 3, Feb. 2018, doi: 10.3389/fnbeh.2018.00003.

[39] E. Butkevičiūtė *et al.*, "Removal of movement artefact for mobile EEG analysis in sports exercises," *IEEE Access*, vol. 7, pp. 7206–7217, 2019, doi: 10.1109/ACCESS.2018.2890335.

[40] A. Kumar, Q. Fang, J. Fu, E. Pirogova, and X. Gu, "Error-related neural responses recorded by electroencephalography during post-stroke rehabilitation movements," *Frontiers Neurorobotics*, vol. 13, p. 107, Dec. 2019, doi: 10.3389/fnbot.2019.00107.

[41] S. Shirinpour, I. Alekseichuk, K. Mantell, and A. Opitz, "Experimental evaluation of methods for real-time EEG phase-specific transcranial magnetic stimulation," *J. Neural Eng.*, vol. 17, no. 4, Jul. 2020, Art. no. 046002, doi: 10.1088/1741-2552/ab9dba.

[42] T.-P. Jung, S. Makeig, A. J. Bell, and T. J. Sejnowski, "Independent component analysis of electroencephalographic and event-related potential data," in *Central Auditory Processing and Neural Modeling*, P. W. F. Poon and J. F. Brugge, Eds. Boston, MA, USA: Springer, 1998, pp. 189–197, doi: 10.1007/978-1-4615-5351-9_17.

[43] T.-P. Jung, S. Makeig, M. J. McKeown, A. J. Bell, T.-W. Lee, and T. J. Sejnowski, "Imaging brain dynamics using independent component analysis," *Proc. IEEE*, vol. 89, no. 7, pp. 1107–1122, Jul. 2001, doi: 10.1109/5.939827.

[44] C.-T. Lin, S.-A. Chen, T.-T. Chiu, H.-Z. Lin, and L.-W. Ko, "Spatial and temporal EEG dynamics of dual-task driving performance," *J. Neuroeng. Rehabil.*, vol. 8, no. 1, p. 11, Dec. 2011, doi: 10.1186/1743-0003-8-11.

[45] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, 1987, doi: 10.1016/0377-0427(87)90125-7.

[46] E. B. Fowlkes and C. L. Mallows, "A method for comparing two hierarchical clusterings," *J. Amer. Statist. Assoc.*, vol. 78, no. 383, pp. 553–569, 1983, doi: 10.2307/2288117.

[47] L. Jäncke, B. Brunner, and M. Esslen, "Brain activation during fast driving in a driving simulator: The role of the lateral prefrontal cortex," *Neuroreport*, vol. 19, no. 11, pp. 1127–1130, Jul. 2008, doi: 10.1097/WNR.0b013e3283056521.

[48] G. Sammer *et al.*, "Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload," *Hum. Brain Mapping*, vol. 28, no. 8, pp. 793–803, Aug. 2007.