IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 16, 2020, accepted January 17, 2020, date of publication January 21, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968360

3D Input Convolutional Neural Networks for P300 Signal Detection

ZEKI ORALHAN
Department of Electrical Electronics Engineering, Nuh Naci Yazgan University, 38090 Kayseri, Turkey
e-mail: zoralhan@nny.edu.tr

ABSTRACT P300 signal is an endogenous event related potential component. It is mostly elicited from the frontal to parietal brain lobes. Electroencephalography is used for acquiring P300 signal from scalp. P300 signal is used for brain-computer interface systems. P300 based brain-computer interface systems are preferable since they have high overall performance. The most significant overall performance indicator is information transfer rate for P300 based brain-computer interface systems. P300 signal detection accuracy and P300 detection time are using for information transfer rate calculation. Hence, P300 signal classification accuracy is important for getting higher information transfer rate. In this study, it is aimed to investigate P300 detection model for higher classification accuracy. Thus, it is proposed 3-dimensional input convolutional neural network model for P300 detection. Moreover, the proposed model was applied with region based P300 speller which constituted audio and visual stimuli. In experiments, the participants were asked to spell desired words in two sessions which were offline and online session. Linear support vector machine, stepwise linear discriminant analysis, 2-dimensional input convolutional neural network, and the proposed method were compared in both online and offline sessions. It is reached highest average classification accuracy rate with the proposed method in both sessions. According to the online session result, average classification accuracy was 94.22% in 3-dimensional input convolutional neural network model. Furthermore, average information transfer rate was 5.53 bit/min in 3-dimensional input convolutional neural network model. We have also applied methods on BCI competition III-dataset II for 2 participants "A" and "B" for evaluating performance of algorithms. The proposed method had higher classification accuracy rate than linear support vector machine, stepwise linear discriminant analysis, 2-dimensional input convolutional neural network, and multi-classifier convolutional neural network which was used in other study on same dataset.

INDEX TERMS Brain computer interface, deep Learning, human machine systems, P300 detection.

I. INTRODUCTION

Brain-computer interface (BCI) systems can be used by healthy or disable people for sending commands to external devices without using neuro-muscular system. Commands are produced via BCI that operated by brain activities [1]. BCI systems can be defined according to the measurement methods of brain activities. There are various BCI systems which based on near infrared spectroscopy (NIRS) [2], [3], functional magnetic resonance (FMRI), electroencephalography (EEG), and magnetoencephalography (MEG) [4], [5]. The most commonly used BCI systems are based on EEG measurement. Because EEG has strong signal response, high signal to noise ratio (SNR), and practical usage [6]. Moreover, EEG measurement is a non-invasive method. Some of EEG based BCI systems use event-related potential (ERP) to produce commands. ERPs are the pattern of brain activities that can be measured by EEG [7]. The P300 is an expectation signal that also can be defined as an endogenous ERP component. There are various paradigm such as single-stimulus, Oddball, and three-stimulus [8]. In BCI systems, Oddball Paradigm is most commonly used in order to elicit P300 wave [9]. According to Oddball Paradigm both standard stimuli with high probability and target stimuli with low probability are presented to user randomly. The subject asked to intend to focus target stimuli. The P300 is elicited 250-500 ms

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

after the occurring target or rare stimuli which are based on auditory, somatosensory or visual [10]. The P300 also can be called P3 since the wave reached its peak point positively in EEG signal around 300 ms after stimuli [11]. The latency of P300 mostly depends on mental performance [12]. P300 signal was discovered in 1965 by Sutton *et al.* [13]. Cognitive fatigue is reduced by using P300 signal in BCI systems. Moreover, signal response is stronger in P300 based BCI systems [14]. The distribution of P300 signal is over the midline of scalp. P300 signal is elicited Fz, Cz, and Pz EEG electrodes position according to international 10-20 system. Typically, P300 signal magnitude increases from the frontal to parietal side [15].

There are significant steps for set up a P300-based BCI system. These steps are: user task, EEG signal recording, signal pre-processing, feature extraction, signal classification, and feedback to user interface [16]. The user task is very important for occurring P300 signal. An efficient user task provides strong P300 signal response, and short latency duration. Also, it protectives eye fatigue [17]. After the EEG recording, in pre-processing step noise reduction and removing the artifacts are provided. Dimensionality reduction process is operated in feature extraction step and classifiers are used for eliciting P300 wave in signal classification steps [18]. Then, it is needed to feedback to user for produced command. Overall performance of a P300 based BCI system is measured by information transfer rate (ITR). P300 detection time and P300 signal classification accuracy are the most important parameters for an enhanced ITR value [19].

Farwell and Donchin designed P300 based-BCI system with using visual stimuli P300 speller for the first time [20]. In 1988 they prepared 36 visual stimuli in 6 × 6 matrix form. Each column and row intensifies randomly. A user of the P300 speller focuses the target visual stimulus which intensifies rarely. In the matrix form P300 speller, each row and column intensify randomly. The user counts him/herself intensifying of target stimuli. Thus, oddball paradigm is occurred [21]. The target character is selected according to in which row and column whether P300 potential is elicited or not. There are many research about visual P300 speller in literature, such as stimuli color, font, size, inter stimulus interval time, and matrix size [22]. But the core structure difference of P300 speller was proposed by Fazel Rezai and Abhari. They prepared a regional based P300 speller. They showed that with using regional P300 speller with higher classification accuracy than row-column based P300 speller [23]. Besides, there are many researches which used only visual stimuli based P300 speller or only auditory stimuli based P300 speller. Schreuder et al. used auditory spelling interface for the occurring P300 wave in their proposed BCI system. They showed that auditory stimuli is enough for BCI system without visual stimuli. They reached averagely 5.26 bits/min ITR value in their experiment. Moreover, the task was complete by 76% of the participants and they used a linear classifier and decision making algorithm [24]. Higashi et al. presented that auditory P300 based brain-computer interface system with using auditory steady state responses (ASSR). They supported the validity of ASSR for BCI systems [25].

In many researches, visual and auditory stimuli were used together in P300-based BCI systems in order to reach higher ITR value, by reducing mental and eye fatigue. Gao et al. reviewed hybrid P300 spellers in literature [26]. Belitski et al. proposed row column based P300 speller with using auditory and visual stimuli together. They carried out the experiments in three different section. Their sections were related on P300 speller with using only visual stimuli, only auditory stimuli, and audio-visual stimuli together. They reached highest classification accuracy with using audio-visual stimuli together [27]. Qu et al. used three-dimensional audio-visual based P300 speller. They compared their proposed three-dimensional speller with two-dimensional speller. Higher classification accuracy was observed in three-dimensional P300 speller [21]. In recent studies such as Zhao et al. contributed body of knowledge with their proposed audio-visual based row column P300 speller. They reached better performance with using hybrid P300 speller [28].

In our experiments we used region based P300 speller with auditory and visual stimuli together since that kind of P300 speller had better performance according to previous researches [23]–[28].

Another significant factor is classification method for higher classification accuracy and ITR value in P300-based BCI systems. Various researches are existing in literature about classification methods in order to reach higher signal classification accuracy. Linear discriminant analysis (LDA) and its derivatives such as Fisher's linear discriminant analysis (FLDA), and stepwise linear discriminant analysis (SWLDA) are mostly preferred for classifying the data in different segmentation [29]. Wu et. al. casted common spatial patterns algorithm in probabilistic modeling setting. The proposed model provided generic EEG spatio-temporal modelling [30]. Support vector machine (SVM) is being used widely in P300 based BCI systems [31]. Other classifiers such as the Bayseian Machine learning [32], Non-linear Bayesian classifiers, and k nearest neighbor (kNN), are also used in P300 detection [33]. But non-linear methods are not preferable in P300 based BCI systems as much as linear classifiers. Borra et al. designed P300 based BCI with using Convolutional Neural Networks (CNNs) for Autistic Spectrum Disorder. They used two learning strategies that were within-session and cross session trainings. They reached 92.27% accuracy value with Cross-session training [34]. Shan et al. applied a simple CNN in order to accurate P300 classification. Character spelling accuracy was improved by up to 19.35% with using CNN model according to traditional machine learning algorithms [35], [36]. Flores et al. compared the CNN model with various machine learning methods. Their results with CNN model outperformed other related works [37]. Liu et al. developed novel a CNN by combining Batch Normalization and Droupout methods. They used their proposed model in P300 signals detection. They showed that their proposed method which named BN3 improved the character detection performance in P300 based BCI systems [38].

In this study we aimed to reach higher ITR value in a P300 based BCI system. For this goal, we proposed a novel 3-dimensional input convolutional neural network (3D input CNN) for P300 detection. Moreover, we used region based P300 speller with auditory and visual stimuli together that we developed before [16]. We carried out our experiments in this scope. We compared linear support vector machine (SVM), SWLDA, 2-dimensional input convolutional neural network (2D input CNN) and 3D input CNN method as a proposed method in both online and offline sessions of experiment. We have reached higher classification accuracy rate and ITR value with the proposed method in both sessions of experiment. Consequently, in this study, we proposed 3D input CNN model for the P300 detection. It was applied for the region based P300 BCI systems. The model was outperfomed other models which were used for comparison. Moreover, we used BCI competition III-dataset II for evaluating the performance of methods. Our proposed method also had higher classification accuracy rate than other methods.

II. MATERIALS AND METHODS

In this study we proposed 3D input CNN model for P300 detection in region based P300 speller with using auditory and visual stimuli. We used SWLDA, SVM, and 2D input CNN for comparing the P300 detection result with the proposed method in experiment.

A. EXPERIMENTAL DESIGN

Eight subjects (five male and three female) that without any chronic disease participated in the experiment. Average age of participants was 21 ± 2. Also, the participants had healthy eyes and ears based on self-report.

We have used regional based P300 speller which was developed by us [16]. In first stage, characters were placed in 5 different regions on the screen. Each region contained 6 different character. Moreover, each region had reference number that was written near the region. While each region were intensifying randomly, the reference numbers was spoken by a male voice. The participant focuses visual and auditory stimuli region which contains target character. The BCI system detected the region which contained P300 signal. In the second stage, the characters that in the detected region were spread 6 different regions on the screen. As in first stage, while each region were intensifying randomly, the reference numbers was spoken by a male voice. The target character was selected by P300 detection. In total, the region based P300 speller presented 30 different characters that are punctuation marks and alphabet letters.

In experiments, we used CleveMed BioRadio device with software of BioCapture in order to record EEG data. The device manufacturer is Great Lakes Neuro Technologies that is an US company.

During the experiments, international 10/20 EEG electrodes placement procedure was used. EEG data was acquired according to unipolar EEG recoding method. Reference electrode was placed on to right earlobe. In the current study, experiments were carried out in a dim place. Sampling rate was selected as 500 Hz in order to EEG data recording.

B. EXPERIMENTAL PROCEDURE

In the current research our developed region based P300 speller was used for eliciting P300 wave [16]. Liquid crystal display (LCD) was used in order to feedback and show P300 speller screen. LCD had 1366×768 pixels. Participants of the experiment were seated 60 cm away from the monitor. Experiments were carried out on a day time and in a dim room. Both visual and auditory stimuli were presented to participants via region based P300 spellers. When the visual stimulus (region) intensified, the reference number of the region was spoken at the same time. Both visual and auditory stimuli occurred for 275 ms randomly synchronously in each region. After the both stimuli, there was 125 ms of silence and non-intensified time that is called inter stimulus interval (ISI) time. After ISI time, both stimuli (visual and auditory) in different region was active. As we mentioned before the P300 speller first stage consisted 5 regions. When each region was intensified one time, a sequence was completed. For the first stage a sequence was 2 seconds. In experiments 8 repeated sequences was a trial. When the first stage of P300 speller trial was completed, LCD was darkened during 3.8 seconds. After this waiting time second stage of P300 speller screen was shown. The P300 speller second stage consisted 6 regions. Each region consisted a character. A sequence duration of second stage was 2.4 seconds. When a trial was finished, a character was chosen in the second stage. Hence, a character chosen time was 39 seconds for the offline session of experiment.

Experiments' road map was shown in Fig. 1. We carried out experiments in two sessions. Firstly, we presented a task in the offline session of the experiments. Each task was spelling a word. Totally 5 tasks were asked to participant to spell. In total, 15 different characters were in tasks. Participants spelled each task two times. We analyzed acquired EEG data with four different signal classification methods. SWLDA, SVM, 2D input CNN, and 3D input CNN models were used in offline and online sessions of experiment. We compared result of classification accuracies according to the methods. Furthermore, the analyzed dataset of offline session were used for training data of online session as shown in Fig. 1. Hence, we acquired 30 characters training dataset for online session of experiments.

After the offline session was completed, one day was break for the experiment. After a day, online session experiment with using SWLDA method was carried out. Than one day was break again. After break, online session experiment with using SVM method was carried out. Also we gave a break day between online session experiments with using 2D input CNN and 3D input CNN models. Thus, each online

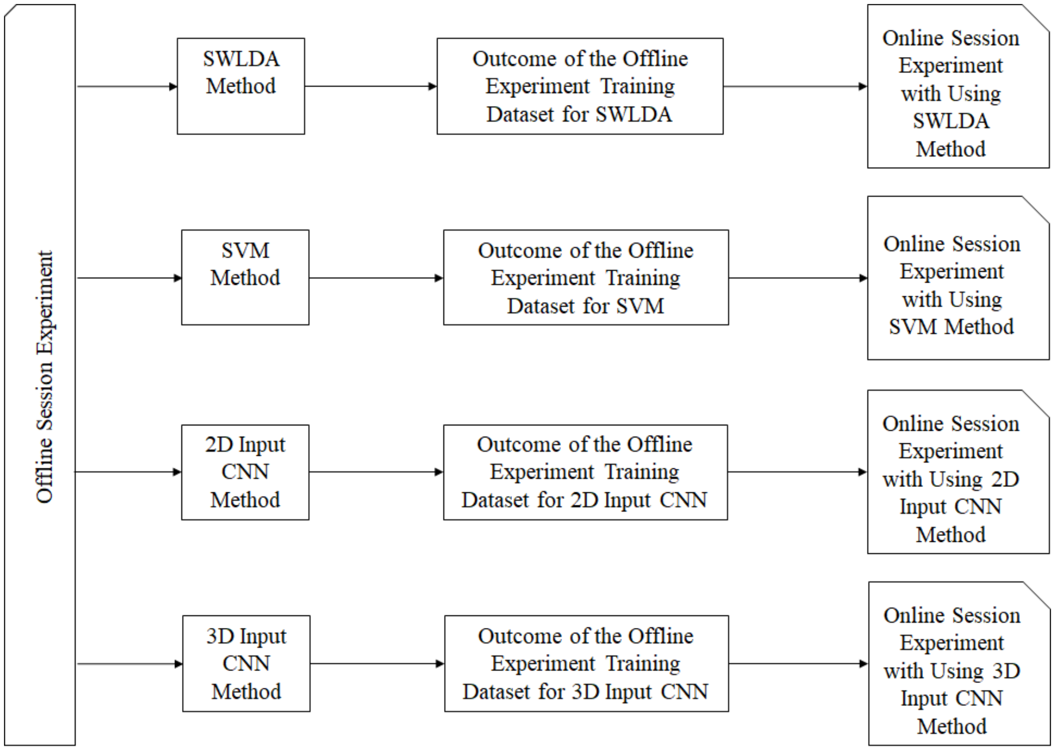

FIGURE 1. Experiments road map.

session experiment was carried out in different days with one day break. Because we wanted to reduce eye fatigue effect and prevent to remember spelling word mentally with a day break between experiments. The experiment participants were asked to spell the word of "POSITIVE" with 7 different characters which were included in training session for 10 times. In online session, when the BCI system detected P300 wave, a trial was interrupted. Thus, a character chosen times was below than 39 seconds for participants.

C. DATA ANALYSIS

We acquired EEG data from 7 EEG electrodes which were placed on Pz, P8, P7, Cz, C4, C3, and Fz EEG channel position. P300 signal was observed dominantly from these positions of EEG electrodes [39]. Thus, we used these channels. Electrodes were placed according to 10/20 international EEG electrodes placement procedure.

In the offline session of experiment we used SWLDA, 3D input CNN, and SVM methods for P300 detection. In the online session of experiment, we used SWLDA and 3D input CNN methods for comparing methods' result. We analyzed EEG data in three steps. In first step, we applied the 5th order Butterworth filter which is band-pass with 0.1 and 35 Hz border frequencies. EEG signal contains eye movements, blinking and muscle movements signals. In the pre-processing step, we applied the filter that provided to eliminate the noise and unwanted artifacts. Moreover, some artifacts can have higher amplitude values than EEG signals. In order to remove these signals, winsorization process was applied on the signal [40]. The highest 10% of EEG data amplitude value in the EEG data from electrodes was replaced amplitude values by the highest value from the remaining EEG data. In the second step, window with 800 ms was started from the stimulus on status [41]. Thus, EEG signals were isolated by window in order to analyze. By using window and 500 Hz sampling rate 400-point data set was elicited. We needed to reduce signal points for classification so with sub-sampling 40-point data set was obtained from 400 point. We crated single feature vector with adding each EEG channel signal points in the same time period. Thus, 280-point data set was acquired by 7 channel EEG data with 40-point dataset.

In third step, SWLDA, 3D input CNN, and SVM were applied separately, in order to the P300 signal classification. SWLDA is the extension of the Fisher's linear discriminant analysis (FLDA) [35]. SWLDA is the most commonly used method for P300 classification. Researches show that classification accuracy with SWLDA is higher in P300 speller based BCI systems [35], [42].

P300 signal classification is a binary problem. The significant point is to observe whether P300 wave is existed in the EEG data. We can define the decision hyper plane as in (1). For our experiment, 280-point data set feature vector is denoted by x, feature vector weight is denoted by w, and bias term is denoted by b in (1).

$$w \cdot x - b = 0. \tag{1}$$

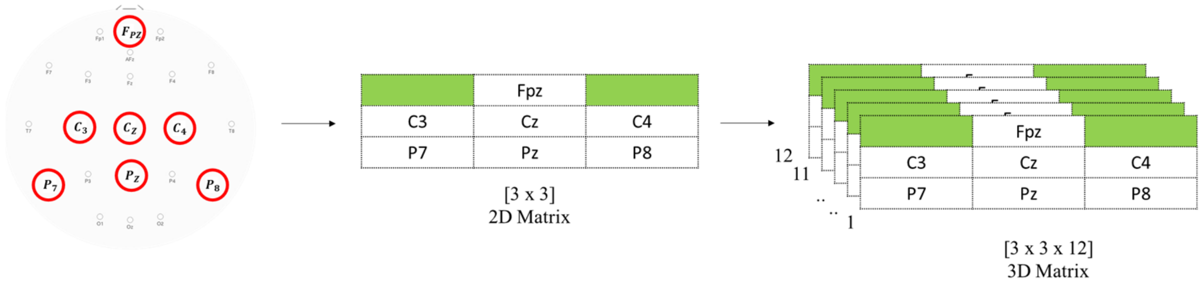

FIGURE 2. The experiment input data construction for 3-dimensional convolutional neural network model: EEG electrodes channel matrix position 1D, 2D, and 3D.

P300 is elicited in order to a region visual intensification and auditory stimulus in our region based P300 speller. P300 signal classification result is calculated by feature vector which maximum of the sum as in (2) and (3).

$$\text{predicted 1st stage} = \max_{1\text{ststage}}\left[\sum_{i_{1\text{ststage}}} w \cdot x_{i_{1\text{ststage}}}\right] \tag{2}$$

$$\text{predicted 2nd stage} = \max_{2\text{ndstage}}\left[\sum_{i_{2\text{ndstage}}} w \cdot x_{i_{2\text{ndstage}}}\right] \tag{3}$$

Predicted region keeps the target characters in the first stage. Character is predicted according to the (3) in second stage. Discriminant function is selected via adding amplitude and channel information to the linear equation. Discriminant function co-efficient is defined by using training dataset. In the current study, the points were selected from the training data set in order to classify the step. P-value was used for selection of estimation.

In first step, the discriminant function did not have any EEG data feature. The feature that has p<0.1 value was added to the discriminant function in each step. Hence, input function was created. But the feature that has p>0.15 value was eliminated by backward stepwise analysis. When the features are not existed to add stepwise analysis ended. Finally, the EEG data was classified according to consisting P300 signal existing.

CNN is based on multilayer perceptron model. Indeed, CNN is a methodology and model for especially image recognition, and related tasks. It is also known that CNN can be implemented to time-series data. EEG is a time-series data thus, we applied CNN model and proposed 3D input CNN model for P300 recognition.

The EEG electrodes were placed on scalp's specific positions where the P300 wave was dominantly existed. In our experiments, in order to prevent information, Seven EEG electrodes position implemented onto a 3×3 matrix as shown in Fig. 2 at the top of this page. Moreover, time information was added via extra axis in order to use CNN operator. Thus, we elicited 3D matrix as 3D input shown in Fig. 2. The green cells are set to zero, since they are not the related on EEG channels. In this current research we used 3D input and generated 12 feature maps as shown in Fig. 2. The CNN are constituted of convolutional layer, pooling layer and connected layer as shown in Fig. 3. Hidden units' local connectivity, pooling and parameter sharing is three significant properties of CNN. A unit that in the hidden layer cannot connected to all units from the previous layer. This property reduces the number of parameters, hence it prevents overfitting. On the other hand, all hidden units were defined by a single feature map and hidden layer has a lot feature maps. A feature map is defined by (4).

$$y^k = f(W_k * x + b_k) \tag{4}$$

where y^k is the 3D array that the k^{th} indices feature map, x is the input as given 3D array, W_k 3D filter that connects k^{th} indices feature map to input, and b_k is the bias term. Sigmoid function is f and convolutional operation is denoted by "$*$" operator. According to the (4), bias term is added to $W_k * x$ array and let W_k has the size $r \times s \times t$. We can define the convolution of x as $[W_k * x](i, j, k)$ in (5).

$$\sum_{u=0}^{r-1}\sum_{v=0}^{s-1}\sum_{w=0}^{t-1} W_k(r-u, s-v, t-w) \times x(i+u, j+v, k+w) \tag{5}$$

Thus, size of input map with the convolution $m \times p \times q$ and a filter size $r \times s \times t$ gives an output size $(m - r + 1) \times (p - s + 1) \times (q - t + 1)$. In our experiment, we applied the convolutions within the all bases, hence 12 feature maps were obtained. The patches size was $2 \times 2 \times 2$, after the convolution size of a feature map was $(3 - 2 + 1) \times (3 - 2 + 1) \times (12 - 2 + 1) = 2 \times 2 \times 11 = 2 \times 2 \times 11$. We have also applied max-pooling operation with $2 \times 2 \times 2$ size in order to reduce feature map size. Thus, size of max pooled feature map $(2 \div 2) \times (2 \div 2) \times (11 \div 2) = 1 \times 1 \times 5$. With 12 feature maps, 60 outputs $(12 \times 1 \times 1 \times 5)$ were elicited. We selected a hidden layer with 20 units via sigmoid function and output layer that consisted 2 units' soft-max function. These 2 units related on conditional probabilities whether P300 wave is existed or not in the input data. Moreover, mini batch gradient descent was used for training the network layer. Hidden layer's weight were initialized randomly, and soft-max layer's weight were initialized to zero.

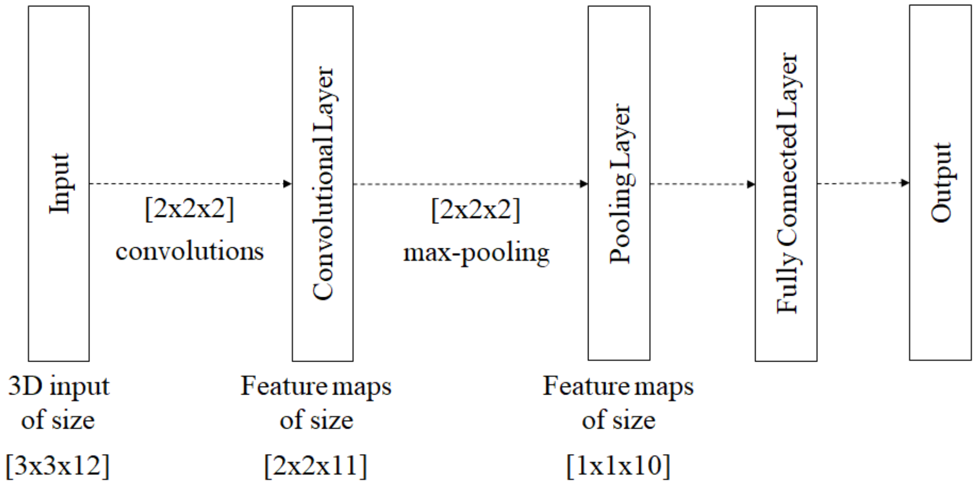

FIGURE 3. Convolutional neural network general design.

SVM was regarded as one of classifiers in P300 based BCI systems [43]. In this study, we used linear SVM in order to find the separating hyperplane. It provides to separate two classes. The distance between the nearest points from classes and the hyperplane is maximal. Hence, margin between both classes is needed to maximize [43]. In our study, we used $\min_{w,b,e}\left(\left(\frac{1}{2}\right)w^T w\right)+\gamma\sum_{i=1}^{N}e_i^2$ to solve problem of minimization according to (6).

$$y_i(w^T f^i + b) = 1 - e_i, \quad i = 1, \ldots, n \tag{6}$$

where f^i denotes the training points and y_i denotes the associated output. This method is also known as linear least squares support vector machine.

Overall BCI system performance was assessed according to information transfer rate (ITR). Equation (7) shows that ITR calculation.

$$\mathrm{ITR} = \frac{60}{T}\left[\log_2 N + P\log_2 P + (1-P)\log_2\left[\frac{1-P}{N-1}\right]\right] \tag{7}$$

According to the (7), signal classification accuracy is denoted by P, total stimulus is denoted by N, and character selection time is denoted by T. In our study, N value was 11, because just regions were intensifying in our region based P300 speller.

III. RESULTS

Offline session of experiment's EEG data were analyzed according to the SWLDA, 3D CNN input, 2D CNN input, and SVM classification methods, respectively. In all methods, P300 signal classification result was recorded for each sequence. As we mentioned before each trial consists 8 repeated sequence. After a trial, classification results were counted. The region which has higher value was selected as a classification result. The same recording process was applied to second stage of P300 speller.

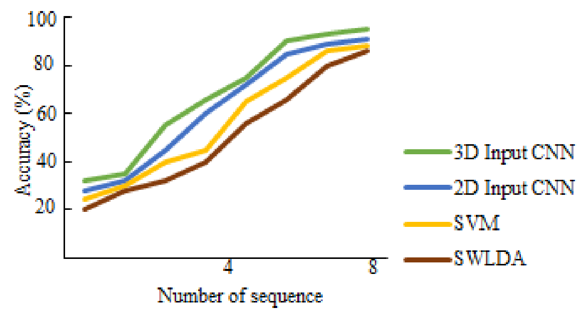

FIGURE 4. Average signal classification result according to 3D Input CNN model, 2D Input CNN model, SVM, and SWLDA with using offline experimental data.

TABLE 1. Average information transfer rate values according to methods with using offline experimental data.

Methods	Information Transfer Rate in bit/min (bpm)
3D CNN Input	4.78
2D CNN Input	4.35
SVM	4.22
SWLDA	3.97

According to offline analysis, we reached highest classification accuracy value 96.25% with 3D CNN input model as shown in Fig. 4. Classification accuracy values were 92.5%, 91.25%, and 88.75% with using 2D CNN input, SVM and SWLDA, respectively.

According to signal classification methods which applied to offline experimental data, information transfer rate values were given in Table 1. We reached highest ITR value as 4.78 bit/min with 3D CNN input model.

In the online session, participants were asked to spell the word of "POSITIVE" for 10 times. Table 2, 3, 4, and 5 show that classification results for each letters according to SWLDA, SVM, 2D input CNN model, and 3D input CNN model, respectively. Moreover, Table 6 consisted all

TABLE 2. Online experiment classification result for each letters with SWLDA method.

	P	O	S	I	T	I	V	E
P	74	7	2	3	3	1		
O	3	68	7	2	1			
S	1	1	65		2	1	11	1
I		1	1	71	1		1	
T			3		65	2	7	
I	2	1			7	73	3	1
V		2	1	3	1	1	58	
E			1	1		2		78

TABLE 3. Online experiment classification result for each letters with SVM method.

	P	O	S	I	T	I	V	E
P	70	2	3			2	2	
O	3	72	2	3	2		4	
S	3		66	1		3	4	
I			1	76	3	2	3	3
T			3		74	2		
I	2		2			69		1
V		4				2	67	
E	2	2	3		1			76

TABLE 4. Online experiment classification result for each letters with 2D input CNN model.

	P	O	S	I	T	I	V	E
P	72	1	3			1	2	
O	3	71	2	3	2		4	2
S	3		69	1		3	4	
I		1		74	3	2	3	3
T			3		74			
I				2		73		
V		4				1	67	
E	2	3	3		1			75

TABLE 5. Online experiment classification result for each letters with 2D input CNN model.

	P	O	S	I	T	I	V	E
P	76	3		1	1		1	
O	2	74	4			1		1
S			73		1		5	
I		1		78				
T		1	2		75		3	
I	1				3	78	1	
V							70	
E	1	1	1	1		1		79

participants' classification accuracy rate result. The average classification accuracies values were 86.25%, 89.06%, 89.84%, and 94.22% in the SWLDA method, SVM method, 2D input CNN model, and 3D input CNN model, respectively. It is understood that the signal classification accuracy in the 3D input CNN model was significantly higher than other methods. Signal classification accuracy in the 3D input CNN Model was improved 9.24%, 5.79%, and 4.88% according to SWLDA, SVM, and 2D input CNN model, respectively.

In the online session of experiment, we observed that the highest average classification accuracy was 98.75% with spelling the letter of "E" in 3D input CNN model.

TABLE 6. Online experiment classification accuracy values for each participants according to the classification methods.

Participants	Classification Accuracy (%)			
	SWLDA	SVM	2D CNN Input	3D CNN Input
P-1	85	88.75	90	93.75
P-2	90	92.5	92.5	96.25
P-3	82.5	85	87.5	91.25
P-4	95	93.75	86.25	96.25
P-5	86.25	91.25	91.25	95
P-6	83.75	86.25	90	92.5
P-7	76.25	81.25	86.25	90
P-8	91.25	93.75	95	98.75
Average	86.25	89.06	89.84	94.22

TABLE 7. Overall performance of online experiment with using SWLDA, SVM, 2D input CNN, and 3d input CNN: information transfer rate values in bit/min (Bpm) for each participant.

Participants	ITR (bit/min)			
	SWLDA	SVM	2D CNN Input	3D CNN Input
P-1	4.28	4.69	4.83	5.3
P-2	5.14	5.47	5.47	6.01
P-3	3.68	3.92	4.17	4.57
P-4	6.22	6.03	5.02	6.42
P-5	4.28	4.84	4.84	5.31
P-6	3.91	4.16	4.56	4.84
P-7	3.05	3.47	3.93	4.31
P-8	5.48	5.83	6.01	6.64
Average	4.55	4.87	4.97	5.53

Each participant classification accuracy rate was given in Table 6. The highest classification accuracy was seen in the participant 8 with 98.75% value in 3D input CNN model. In SWLDA method, participant 7 has 76.25% classification accuracy value which was the lowest score in the experiment.

Table 7 shows that the information transfer value in bit/min (bpm). Average ITR value 4.55, 4.87, 4.97, and 5.53 bpm in SWLDA, SVM, 2D input CNN and 3D Input CNN, respectively. The highest ITR value was 6.64 bpm with participant 8 in 3D input CNN model. Besides, ITR is the most important parameter for the BCI overall performance. We reached the highest average ITR value with the proposed method.

IV. DISCUSSION

In this current study, we wanted to research whether our proposed model outperform compared to other methods which were widely using in P300 detection or not. Besides, in literature there are some researches that compared the CNN models between machine learning algorithms for P300 detection [34]–[37]. Moreover Liu et. al developed enhanced CNN model which named BN3 [38]. They outperformed with their proposed method than CNN and SVM. We realized that in P300 based BCI systems classification accuracy can reach higher values with enhanced CNN models. Thus, we proposed 3D input CNN method for P300 detection.

We compared the proposed method with SWLDA, SVM and 2D input CNN in both online session and offline session

TABLE 8. Performance of methods used in this study evaluated on BCI competition III-dataset II for "A" and "B" participants.

Methods	Classification Accuracy (%) in 5th Sequence			Classification Accuracy (%) in 15th Sequence		
	P-A	P-B	Avg.	P-A	P-B	Avg.
SWLDA	63	72	67.5	82	88	85
SVM	70	80	75	91	94	92.5
2D Input CNN	73	84	77.5	93	96	94.5
3D Input CNN	74	86	80	96	98	97

experiments. The obtained results showed that 3D input CNN outperformed SWLDA, SVM, and 2D input CNN.

We have also applied methods on BCI competition III-dataset II for 2 participants "A" and "B" for evaluating performance of algorithms. Table 8 depicts the classification accuracy results as methods' performance with respect to the 5th and 15th sequences. Classification accuracy value with our proposed method was significantly higher than other methods.

Cecotti and Graser also applied their models on BCI competition III-dataset II. They reached average classification accuracy rate 95.5% with multi-classifier convolutional neural network as highest result [44]. On same dataset, the best average result is obtained with our proposed method with a classification accuracy rate of 97% as shown in Table 8. Both researches' results were acquired according to 15th sequence of intensifying. Our proposed approach both helps to reduce dimensionality and extract hidden features efficiently. That's why the proposed method outperformed compared to others.

As we mentioned before, in our study we used offline session experiment outcome data as training dataset for online session of experiment. As a recent study, Jin et. al. introduced generic model set. They used 116 subjects' ERP data to train their generic model set. Thus, they obtained ten models that were trained by weighted linear discriminant analysis. When new subjects join the experiment, their result showed that new subjects matched the suitable generic model [45]. This method can be used for increasing performance and reducing calibration time in our future studies.

Moreover, Chaurasiya et. al. proposed weighted edit distance based P300 speller [46]. Jin et. al. compared the mismatch negativity paradigm with single character presentation approach. They reached higher classification accuracy rate and ITR value. N200 and N400 components were elicited larger with Jin et. al. approach [47]. Our proposed signal classification method can be applied on dataset which was elicited from mismatch negativity paradigm and weighted edit distance approach. Thus, higher classification accuracy and ITR value can be observed.

For further studies, we will also asset our result with other artificial neural networks [38], [48].

V. CONCLUSION

The current research investigated whether it is possible to use 3D input CNN model for P300 detection or not. Also, we compared the SWLDA method, SVM and 3D input CNN model for P300 detection. According to our experiment we improved average classification accuracy with 3D input CNN model. Moreover, it is the first time that 3D input CNN model was applied to region based P300 speller. We conducted an Analysis of variance (ANOVA) to assets. According to ANOVA there was significant difference in classification accuracy rate between methods ($F = 4.785$, $p = 0.008$). According to ANOVA result we used post-hoc least significant difference (LSD) test to evaluate classification accuracy rate. There was statistically significant difference in classification accuracy rate between using 3D input CNN method and other methods. (SWLDA $p = 0.001$; SVM $p = 0.022$; 2D input CNN $p = 0.05$).

REFERENCES

[1] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain–computer interface spellers: A review," *Brain Sci.*, vol. 8, no. 4, p. 57, Mar. 2018.

[2] F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. M. Pavia, U. Wolf, and M. Wolf, "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," *NeuroImage*, vol. 85, pp. 6–27, Jan. 2014.

[3] S. Zhang, Y. Zheng, D. Wang, L. Wang, J. Ma, J. Zhang, W. Xu, D. Li, and D. Zhang, "Application of a common spatial pattern-based algorithm for an fNIRS-based motor imagery brain–computer interface," *Neurosci. Lett.*, vol. 655, pp. 35–40, Aug. 2017.

[4] B. He and Z. Liu, "Multimodal functional neuroimaging: Integrating functional MRI and EEG/MEG," *IEEE Rev. Biomed. Eng.*, vol. 1, pp. 23–40, 2008.

[5] J. Del R. Milan and J. Carmena, "Invasive or noninvasive: Understanding brain-machine interface technology conversations in BME," *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 1, pp. 16–22, Jan. 2010.

[6] Z. Oralhan and M. Tokmakçi, "The effect of duty cycle and brightness variation of visual stimuli on SSVEP in brain computer interface systems," *IETE J. Res.*, vol. 62, no. 6, pp. 795–803, Nov. 2016.

[7] Z. Oralhan, "Advanced SSVEP stimulator for brain–computer interface and signal classification with using convolutional neural network," *Electron. Lett.*, vol. 55, no. 25, pp. 1329–1331, Dec. 2019.

[8] M. M. Walsh, G. Gunzelmann, and J. R. Anderson, "Relationship of P3b single-trial latencies and response times in one two and three-stimulus oddball tasks," *Biol. Psychol.*, vol. 123, pp. 47–61, Feb. 2017.

[9] J. Jin, B. Z. Allison, Y. Zhang, X. Wang, and A. Cichocki, "An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions," *Int. J. Neur. Syst.*, vol. 24, no. 8, Dec. 2014, Art. no. 1450027.

[10] Z. Oralhan, "2 stages-region-based P300 speller in brain–computer interface," *IETE J. Res.*, vol. 65, no. 6, pp. 740–748, 2019.

[11] J. Polich, "Updating P300: An integrative theory of P3a and P3b," *Clin. Neurophysiol.*, vol. 118, no. 10, pp. 2128–2148, Oct. 2007.

[12] E. Saliasi, L. Geerligs, M. M. Lorist, and N. M. Maurits, "The relationship between P3 amplitude and working memory performance differs in young and older adults," *PLoS ONE*, vol. 8, no. 5, May 2013, Art. no. e63701.

[13] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, no. 3700, pp. 1187–1188, Nov. 1965.

[14] G. Townsend, J. Shanahan, D. B. Ryan, and E. W. Sellers, "A general P300 brain–computer interface presentation paradigm based on performance guided constraints," *Neurosci. Lett.*, vol. 531, no. 2, pp. 63–68, Dec. 2012.

[15] L. Li, C. Gratton, D. Yao, and R. T. Knight, "Role of frontal and parietal cortices in the control of bottom-up and top-down attention in humans," *Brain Res.*, vol. 1344, pp. 173–184, Jul. 2010.

[16] Z. Oralhan, "A new paradigm for region-based P300 speller in brain computer interface," *IEEE Access*, vol. 7, pp. 106618–106627, 2019.

[17] Y. Li, C. S. Nam, B. B. Shadden, and S. L. Johnson, "A P300-based brain–computer interface: Effects of interface type and screen size," *Int. J. Hum.-Comput. Interact.*, vol. 27, no. 1, pp. 52–68, Dec. 2010.

[18] Q. Wu, Y. Zhang, J. Liu, J. Sun, A. Cichocki, and F. Gao, "Regularized group sparse discriminant analysis for P300-based brain-computer interface," *Int. J. Neural Syst.*, vol. 26, no. 6, 2019, Art. no. 1950002.

[19] Y. Lu and L. Bi, "EEG signals-based longitudinal control system for a brain-controlled vehicle," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 2, pp. 323–332, Feb. 2019.
[20] L. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.
[21] J. Qu, F. Wang, Z. Xia, T. Yu, J. Xiao, Z. Yu, Z. Gu, and Y. Li, "A novel three-dimensional P300 speller based on stereo visual stimuli," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 4, pp. 392–399, Aug. 2018.
[22] M. Guo, J. Jin, Y. Jiao, X. Wang, and A. Cichocki, "Investigation of visual stimulus with various colors and the layout for the oddball paradigm in ERP-based BCI," *Frontiers Comput. Neurosci.*, vol. 13, Apr. 2019, Art. no. 00024.
[23] R. Fazel-Rezai and K. Abhari, "A region-based P300 speller for brain-computer interface," *Can. J. Elect. Comput. Eng.*, vol. 34, no. 3, pp. 81–85, 2009.
[24] M. Schreuder, T. Rost, and M. Tangermann, "Listen you are writing! Speeding up online spelling with a dynamic auditory BCI," *Front. Neurosci.*, vol. 5, 2011, Art. no. 112.
[25] H. Higashi, T. M. Rutkowski, Y. Washizawa, A. Cichocki, and T. Tanaka, "EEG auditory steady state responses classification for the novel BCI," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Boston, MA, USA, Aug. 2011, pp. 4576–4579.
[26] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.
[27] A. Belitski, J. Farquhar, and P. Desain, "P300 audio-visual speller," *J. Neural Eng.*, vol. 8, no. 2, Apr. 2011, Art. no. 025022.
[28] X. Zhao, G. Cui, L. Yuan, T. Tanaka, Q. Zhao, and J. Cao, "A hybrid brain computer interface based on audiovisual stimuli P300," in *Proc. IEEE Int. Conf. Consum. Electron.-Asia (ICCE-Asia)*, Jeju, South Korea, Jun. 2018, pp. 206–212.
[29] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. Mcfarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 Speller," *J. Neural Eng.*, vol. 3, no. 4, pp. 299–305, Dec. 2006.
[30] W. Wu, Z. Chen, X. Gao, Y. Li, E. N. Brown, and S. Gao, "Probabilistic common spatial patterns for multichannel EEG analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 639–653, Mar. 2015.
[31] T. Lal, M. Schroder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer, and B. Scholkopf, "Support vector channel selection in BCI," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1003–1010, Jun. 2004.
[32] W. Wu, S. Nagarajan, and Z. Chen, "Bayesian machine learning: EEG/MEG signal processing measurements," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 14–36, 2016.
[33] A. Bablani, D. R. Edla, and S. Dodia, "Classification of EEG data using k-nearest neighbor approach for concealed information test," *Procedia Comput. Sci.*, vol. 143, pp. 242–249, Jan. 2018.
[34] D. Borra, S. Fantozzi, and E. Magosso, "Convolutional neural network for a P300 brain-computer interface to improve social attention in autistic spectrum disorder," in *Proc. 15th Medit. Conf. Med. Biol. Eng. Comput.*, Coimbra, Portugal, vol. 76, 2019, pp. 206–212.
[35] H. Shan, Y. Liu, and T. Stefanov, "A simple convolutional neural network for accurate P300 detection and character spelling in brain computer interface," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Stockholm, Sweden, Jul. 2018, pp. 1604–1610.
[36] H. Shan, Y. Liu, and T. Stefanov, "Ensemble of convolutional neural networks for P300 speller in brain computer interface," in *Proc. Int. Conf. Artif. Neural Netw.*, Munich, Germany, 2019, pp. 376–394.
[37] C. Flores, V. Flores, D. Achanccaray, and J. Andreu-Perez, "A convolutional neural network approach for a P300-based brain-computer interface for disabled and healthy subjects," in *Proc. 10th Comput. Sci. Electron. Eng. (CEEC)*, Colchester, Britain, Sep. 2018, pp. 192–197.
[38] M. Liu, W. Wu, Z. Gu, Z. Yu, F. Qi, and Y. Li, "Deep learning based on Batch Normalization for P300 signal detection," *Neurocomputing*, vol. 275, pp. 288–297, Jan. 2018.
[39] V. Changoluisa, P. Varona, and F. B. Rodriguez, "An electrode selection approach in P300-based BCIs to address inter- and intra-subject variability," in *Proc. 6th Int. Conf. Brain-Comput. Interface (BCI)*, GangWon, South Korea, Jan. 2018, pp. 1–4, doi: 10.1109/iww-bci.2018.8311497.
[40] M. Fabiani, G. Gratton, D. Karis, and E. Donchin, "Definition identification and reliability of measurement of the P300 component of the event-related brain potential," *Adv. Psychophysiol.*, vol. 2, no. 3, pp. 1–78, 1987.
[41] T. Castermans, M. Duvinage, M. Petieau, T. Hoellinger, C. D. Saedeleer, K. Seetharaman, A. Bengoetxea, G. Cheron, and T. Dutoit, "Optimizing the performances of a P300-based brain–computer interface in ambulatory conditions," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 1, no. 4, pp. 566–577, Dec. 2011.
[42] D. Klobassa, T. Vaughan, P. Brunner, N. Schwartz, J. Wolpaw, C. Neuper, and E. Sellers, "Toward a high-throughput auditory P300-based brain–computer interface," *Clin. Neurophysiol.*, vol. 120, no. 7, pp. 1252–1261, Jul. 2009.
[43] M. Thulasidas, C. Guan, and J. Wu, "Robust classification of EEG signal for brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 1, pp. 24–29, Mar. 2006.
[44] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain-computer interfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 433–445, Mar. 2011.
[45] J. Jin, S. Li, I. Daly, Y. Miao, C. Liu, X. Wang, and A. Cichocki, "The study of generic model set for reducing calibration time in P300-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 3–12, Jan. 2020, doi: 10.1109/tnsre.2019.2956488.
[46] R. K. Chaurasiya, N. D. Londhe, and S. Ghosh, "A novel weighted edit distance-based spelling correction approach for improving the reliability of Devanagari script-based P300 speller system," *IEEE Access*, vol. 4, pp. 8184–8198, 2016.
[47] J. Jin, E. W. Sellers, S. Zhou, Y. Zhang, X. Wang, and A. Cichocki, "A P300 brain–computer interface based on a modification of the mismatch negativity paradigm," *Int. J. Neur. Syst.*, vol. 25, no. 03, May 2015, Art. no. 1550011.
[48] R. Sánchez-Reolid, A. García, M. Vicente-Querol, L. Fernández-Aguilar, M. López, and A. González, "Artificial neural networks to assess emotional states from brain-computer interface," *Electronics*, vol. 7, no. 12, p. 384, Dec. 2018.

ZEKI ORALHAN was born in Turkey. He received the B.Eng. and Ph.D. degrees in electrical electronics engineering from Erciyes University, Kayseri, Turkey, in 2008 and 2016, respectively. From 2008 to 2014, he was a Senior Engineer in a high-tech firm. From 2014 to 2017, he was a Research and a Development Advisor of Türk Telekom. Since 2018, he has been an Assistant Professor with the Electrical Electronics Engineering Department, Nuh Naci Yazgan University, where he has also been a Technology Firm with the Techno Park Center. His research interests include brain computer interface, artificial learning, data mining, and signal processing.