

Light-Weight 1-D Convolutional Neural Network Architecture for Mental Task Identification and Classification Based on Single-Channel EEG

Manali Saini, Udit Satija, *Member, IEEE*, and Madhur Deo Upadhayay, *Member, IEEE*

Abstract—Mental task identification and classification using single/limited channel(s) electroencephalogram (EEG) signals in real-time play an important role in the design of portable brain-computer interface (BCI) and neurofeedback (NFB) systems. However, the real-time recorded EEG signals are often contaminated with noises such as ocular artifacts (OAs) and muscle artifacts (MAs), which deteriorate the hand-crafted features extracted from EEG signal, resulting in inadequate identification and classification of mental tasks. Therefore, we investigate the use of recent deep learning techniques which do not require any manual feature extraction or artifact suppression step. In this paper, we propose a light-weight one-dimensional convolutional neural network (1D-CNN) architecture for mental task identification and classification. The robustness of the proposed architecture is evaluated using artifact-free and artifact-contaminated EEG signals taken from two publicly available databases (i.e., Keirn and Aunon (K) database and EEGMAT (E) database) and in-house (R) database recorded using single-channel neurosky mindwave mobile 2 (MWM2) EEG headset in performing not only mental/non-mental binary task classification but also different mental/mental multi-tasks classification. Evaluation results demonstrate that the proposed architecture achieves the highest subject-independent classification accuracy of 99.7% and 100% for multi-class classification and pair-wise mental tasks classification respectively in database K . Further, the proposed architecture achieves subject-independent classification accuracy of 99% and 98% in database E and the recorded database R respectively. Comparative performance analysis demonstrates that the proposed architecture outperforms existing approaches not only in terms of classification accuracy but also in robustness against artifacts.

Index Terms—Electroencephalogram, Mental task identification, Classification, Deep learning, Convolutional neural network.

I. INTRODUCTION

Electroencephalogram (EEG) represents the electrical activity of the brain [1]. Due to low cost, high temporal resolution and non-invasiveness, EEG is the most commonly used signal in designing neurofeedback (NFB), neural control interface (NCI) and brain computer interface (BCI) systems [2]–[5]. Since portability is one of the critical features for unsupervised mental health monitoring, these systems demand accurate detection of neuronal activities using single/limited channel(s) EEGs [6]. It has been shown that EEG signal exhibits different neuronal changes due to various mental activities including, mental tasks and mental stress [6]–[9].

Manali Saini and Madhur Deo Upadhayay are with the Department of Electrical Engineering, Shiv Nadar University, Greater Noida, Udit Satija is with Department of Electrical Engineering, Indian Institute of Technology Patna, Bihar, India. E-mail: ms905@snu.edu.in, udit@iitp.ac.in, madhur_deo@yahoo.com.

These changes are induced when the subjects are presented with standardized stress tests, workload tasks, and questionnaires by psychologists [7], [8]. Accurate analysis of these neuronal changes enables identification and classification of different mental tasks which is useful for patients suffering from motor, cerebral, and behavioral disorders, for example, attention deficit hyperactivity disorder (ADHD) and autism [10], [11], as well as for healthy persons to improve their concentration and cognitive performance [5]. Furthermore, the identification and classification of mental tasks from EEG are beneficial for early detection of mental stress and diagnosis of several diseases such as, depression, heart attack, etc [12].

Presently, existing techniques exploit the use of various feature extraction techniques and different machine learning classifiers for mental task identification and classification [7], [13]–[16]. Since single/limited channel(s) EEGs are commonly corrupted with various ocular and muscle artifacts, performance of the hand-crafted features-based mental task identification techniques deteriorates significantly [17], [18]. Recently, deep convolutional neural network (CNN) has gained attention due to its ability to extract high level features automatically from the raw data for accurate analysis of different physiological signals [3], [19], [20]. Although CNN has been applied on EEG signals for mental workload level classification, there exists no work which utilizes the efficacy of CNN for mental task identification and classification. Furthermore, existing CNN-based mental workload technique [21] uses time-frequency representation of EEG in 2D form which demands a complex architecture for learning its 2D features and increases the computational complexity [16], [22]. However, real-time NFB system demands low latency in classification process in order to provide timely feedback to the user.

A. Related work and motivation

Numerous works have been reported in the literature for the identification and classification of different types of mental tasks from EEG [13], [14], [19]. In [6], Keirn et al. proposed the use of autoregressive (AR) parameters and band power asymmetry ratios to classify five mental tasks from EEG recordings of five subjects, using Bayes quadratic classifier (BQC). Similar features have been used to train elman neural network (ENN) with resilient backpropagation (RBP) [2], and feed forward neural network (NN) [23], for classification of different mental tasks. In [14], S. Dutta et al., proposed multivariate AR model based features extracted from intrinsic

mode functions in multivariate empirical mode decomposition (MEMD) domain, to classify three cognitive tasks using least squares support vector machine (LS-SVM) classifier. In [15], modified lempel–Ziv (LZ) complexity has been presented along with band powers and entropy as features to discriminate five mental tasks using K-nearest neighbour (K-NN) and linear discriminant analysis (LDA) classifiers. Power spectral density based features have been fed to LDA classifier for classification of five mental tasks in six subjects [4]. In [24], PSD features have also been used along with improved particle swarm optimization (IPSO) based NN classifier to distinguish three mental tasks. Similar features along with statistical features, frequency-bands' power and Higuchi's fractal dimension have been fed to SVM for classification of mental arithmetic tasks in ten subjects [5]. In [25], Alyasseri et al. used subject-independent discrete wavelet transform (DWT) based statistical features along with entropy to classify five mental tasks for seven subjects using artificial neural network (ANN). In [26], EEG signals recorded from 41 subjects during three mental tasks have been classified using subject-independent statistical features and multi-layer perceptron (MLP) kernel based SVM. An immune-feature weighted SVM has been proposed to classify five mental tasks for seven subjects with approximate entropy feature in [27].

In [19], the EEG waves obtained from discrete wavelet transform of the artifact-removed signal are used as inputs to a 1-D CNN model for discriminating different levels of multimedia learning tasks in 34 subjects. In [7], Z. Pei et al. utilized EEG features exhibiting intra-channel and inter-channel information to classify multiple workload tasks with an overall accuracy of 85%. In [28], P. Zhang, et al., proposed a recurrent three dimensional (3-D) CNN to classify high and low mental workload levels across two tasks, i.e., spatial n-back task and an arithmetic task for twenty subjects. A custom domain adaptation based 3-D CNN with the spatial, spectral and temporal inputs has been used to distinguish four levels of cognitive load from 13 subjects in [3]. In [21], Jiao et al., proposed a single-channel spectral- and temporal-spectral-map based CNN model to classify four different levels (low to high) of working memory while recalling some characters from a set shown to 15 participants.

Most of the existing techniques use subject-dependent hand-crafted features and conventional machine learning approaches for mental task identification and classification [3]. However, these techniques may not generalize across subjects and databases due to high structural and functional changeability between subjects and the non-stationarity of EEG [3]. This issue can be addressed by the use of deep learning approaches where the features are automatically learnt from the raw data during the training process [29]. One of the most popular deep learning approaches is CNN, which has been successfully applied for various classification problems related to EEG signals including, seizure detection, emotion recognition, and mental workload level classification [21], [28]–[32]. However, most of these works utilize artifact removal preprocessing step to improve classification accuracy [29] and/or time-frequency representation of EEG signal as a 2-D or 3-D input to CNN which increases the computational complexity

of the complete system [28], [30]–[32]. Though, there exists two CNN-based subject-independent mental task classification techniques that classify resting state from multimedia learning task and different mental tasks respectively, they use artifact removal step and DWT representation of the EEG signal [19] and achieve low accuracy (70%) [33]. Also, the overall CNN architecture becomes computationally complex. Therefore, in this paper, we present a light-weight 1-D CNN architecture which involves few layers and extracts features automatically from the raw EEG signals for accurate identification and classification of mental tasks.

B. Objective and key contributions

Literature studies demonstrate that most of the existing hand-crafted feature-based mental task classification techniques have poor subject-independent classification accuracy for both artifact-free and artifact-contaminated EEG signals. Furthermore, use of artifact removal step can alter clinical features of EEG signals even in case of artifact-free EEG signals [17], [18]. Existing CNN-based mental task and mental workload classification techniques use complex architecture and input signal in the form of 2D or 3D time-frequency representations of single/multi-channel EEGs [16]. Therefore, in this paper, we propose a light-weight 1D-CNN architecture for identification and classification of mental tasks from single-channel EEG signal. The significant contributions of this paper are summarized as follows:

- Investigation of 1D-CNN for automated meaningful feature extraction from the raw single-channel EEG signal.
- Proposed a low complex subject-independent architecture based on 1D CNN, using only a few layers.
- Recording of single channel EEG signal using Neurosky mindwave mobile 2 device to examine the performance of the proposed architecture on real-time EEGs.
- Examination of the classification accuracy of the proposed architecture for not only mental/non-mental task classification but also several mental/mental multi-tasks classification unlike existing techniques.
- Examination of the mental task classification accuracy of the proposed architecture using OA- and MA-contaminated EEG signals.
- Evaluation of the proposed architecture on three databases, including two publicly available databases and one real-time recorded database.

The rest of the paper is organized as follows: Section II describes the databases used in this work. Section III describes the major constituents of the proposed architecture. Section IV discusses the performance results obtained on different EEG signals taken from publicly available databases as well as real-time recorded data. Section V concludes the paper.

II. DESCRIPTION OF DATABASES

This section presents a brief description of the three EEG databases used in this work, including two publicly available databases, i.e., Keirn and Aunon database, EEG during mental arithmetic tasks (EEGMAT) database and one recorded database.

A. Keirn and Aunon database (K)

This database was recorded by Keirn and Aunon from seven subjects while performing four mental tasks and one baseline/non-mental task [6]. It consists of both artifact-free and OA-corrupted EEG signals recorded at a sampling rate of $250Hz$ from six electrode positions, i.e., C_3, C_4, P_3, P_4, O_1 and O_2 according to the 10 – 20 system, with A_1 and A_2 as the reference electrodes. Eye blinks (OAs) were also recorded by a separate channel. For each task, recording procedure was performed for 10 seconds in a sound-controlled booth with dim lighting. Different number of trials were performed by each subject for each task. For example, subjects 1, 3, 4 and 6 performed 10 trials, subjects 2 and 7 performed 5 trials and subject 5 performed 15 trials of each task. In this work, we have used all subjects of this database except subject 4 due to incomplete information available for that subject. The database consists of the signals recorded during the following tasks [6]:

- Baseline task (BT): The subjects were informed to relax and sit idle. No task was performed and the data was recorded in ‘eyes closed’ and ‘eyes open’ position.
- Multiplication task (MT): The subjects were given a complex multiplication problem to solve without speaking and making any movements.
- Letter composing task (LT): The subjects were asked to mentally compose a letter to a known person without speaking or making any movement.
- Geometric figure rotation task (RT): The subjects were presented with a complex 3 – D figure for 30 seconds, after which the figure was removed, and they were asked to visualize the same figure being rotated about an axis. The data was recorded during this visualization process.
- Visual counting task (VT): The subjects were asked to visualize some numbers written in a sequential order on a blackboard, with the previous number being removed before the next number was written.

B. EEG during mental arithmetic tasks (EEGMAT) database (E)

Database E consists of EEG signals collected from 36 subjects before and during performance of a mental arithmetic task using Neurocom 23 channel device [34], [35]. It consists of artifact-free EEG signals of 60s duration each, recorded from $F_{P1}, F_{P2}, F_3, F_4, F_z, F_7, F_8, C_3, C_4, C_z, P_3, P_4, P_z, O_1, O_2, T_3, T_4, T_5,$ and T_6 electrodes positioned according to the 10 – 20 electrode placement system. The sampling frequency was kept at $500Hz$. Only one trial with 19 EEG signals was recorded per subject and task. The tasks are as follows:

- No mental task/ baseline task (BT): The subjects did not perform any mental task and were asked to sit in a relaxed position.
- Serial subtraction task (ST): Subjects were instructed to perform a serial subtraction task including 4 digit minuend and 2 digit subtrahend and communicate the results orally. At the end, the number of subtractions were noted based on the communicated results. A good or a bad quality count was given to each subject based on the results.

TABLE I: Database R record details.

Subject number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Gender	M	M	M	F	F	F	F	M	M	F	F	F	M	M	F	M	M	F	M	M
Age (Years)	25	22	27	28	31	27	30	55	25	28	28	27	25	27	30	32	27	29	26	26
Subtractions (number)	8	6	7	9	6	5	7	9	8	9	10	9	10	7	8	7	9	8	7	10

C. Recorded database (R)

To evaluate the effectiveness of the feasibility of single-channel EEG data for mental task identification and classification, we recorded in-house EEG signals using twenty subjects during baseline and mental task activity. Details of these subjects have been described in Table I. Neurosky mindwave mobile 2 (MWM2) headset was used to record single-channel EEG from F_{P1} position before and during the performance of mental arithmetic task. The headset consists of three electrodes, one for EEG (F_{P1}) and other two electrodes for ground and reference (A_1 position) [36], as shown in the recording set up in Fig. 1. EEG acquisition has been performed in a sound controlled laboratory with normal lighting. The inbuilt Thinkgear ASIC (application-specific integrated circuit) module (TGAM) pre-processes the raw signal, i.e., removal of powerline interference (PLI) and MAs. Communication is established between the device and computer using a bluetooth module [36]. The data was recorded at a sampling frequency of $512Hz$ with 12 bit resolution and analyzed in MATLAB software. Five number of trials of mental and baseline tasks were recorded for each subject and each trial lasted for 10 seconds. Trials are the different sessions of EEG signal recording which were performed during the verbal announcement of another person to ‘start’ and ‘stop’ the session. The following tasks were performed:

- Baseline task (BT): Subjects were asked to sit in a relaxed position without making any movement, with ‘eyes open’ and ‘eyes closed’ positions. During this time, the data was labeled as baseline task. This procedure was repeated for all subjects five times, resulting in total five trials/sessions of baseline task per subject.
- Serial subtraction task (ST): Subjects were instructed to perform serial subtraction between one 4 digit number (minuend) and other 2 digit number (subtrahend) without speaking and making any movement, in ‘eyes open’ and ‘closed’ position. After the announcement of ‘start’, participants started performing serial subtraction and communicated their subtraction results after the ‘stop’ announcement. Based on their outcome, the number of subtractions performed by each subject was noted by the person. In each trial, different sets of numbers were given for performing ST. To illustrate an example of ST, let the minuend be 4373 and subtrahend be 59, then the result after first subtraction: $4373 - 59 = 4314$, after second subtraction: $4373 - 59 - 59 = 4314 - 59 = 4254$ and so on. Each participant performed varying number of subtractions depending upon his/her calculation speed.

Fig. 2 shows two EEG signals recorded from 3rd subject in ‘eyes closed’ position while performing serial subtraction and baseline tasks. The difference between the amplitudes

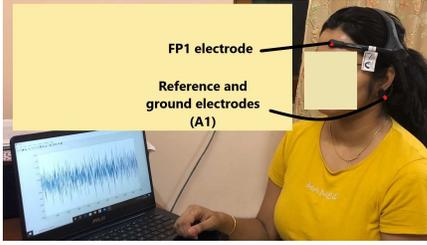


Fig. 1: Recording of EEG data of 4th subject from neurosky MWM2 headset while performing the mental task (ST) in ‘eyes closed’ position.

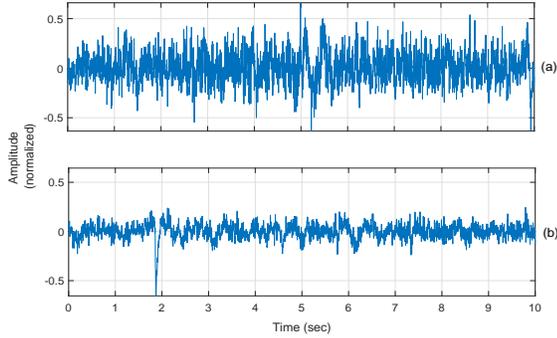


Fig. 2: Recorded EEG signals of 3rd subject (a) and (b) while performing the BT and ST tasks in ‘eyes closed’ position respectively.

and frequencies of the two signals is quite observable from the figure. Since the EEG signals have been recorded at different sampling frequencies for all the three databases, all signals have been re-sampled to 500Hz using cubic spline algorithm [37], [38] for adequate classification. In this work, we assume that the EEG data is always available and it may be corrupted by in-band ocular artifacts and muscle artifacts, baseline wander and powerline interference. Although use of basic pre-processing is essential to eliminate flat line, instrumentation noise or raw noise, for which there are well-established techniques, it is out of scope of this work.

III. PROPOSED 1D-CNN ARCHITECTURE

CNN is a popular deep learning approach that has been successfully applied to EEG signal analysis [28], [39]. It possesses a hierarchical structure in the form of different layers, in which each layer with a specific operation extracts high-level features from the raw input data [40]. In comparison with the conventional fully connected networks, CNN has a tremendous learning capability in extracting more robust features at each layer as well as a good generalization performance [40]. This section presents the proposed 1D-CNN architecture with an input raw single-channel EEG signal denoted as $x[n]$ which is illustrated in Fig. 3. It comprises of two 1D-convolution layers, one 1D-max pooling layer, one flatten layer with dropout and a final dense/ fully connected layer with softmax/ sigmoid activation for classification output.

A. 1D-convolution layer (Conv1D)

The raw one-dimensional EEG signal (vector) $x[n]$, where, $n = 1, 2, \dots, N$, is given as an input to the first layer of the CNN architecture, i.e., conv1D, as shown in the Fig. 3. The layer utilizes the following parameters:

- **Filters / Kernels:** The filters / kernels produce feature maps by performing convolution with the input signal. The number and size of kernels are crucial for adequately capturing relevant features from the signal. Let $k[n]$ denote the convolution kernel with size v , then the convolution output $c[n]$ can be given as:

$$c[n] = x[n] * k[n] = \sum_{m=0}^{v-1} x[m] \cdot k[n-m] \quad (1)$$

where, ‘*’ denotes the convolution operation. In general, the convolved feature at the output of l^{th} layer can be written as [41]:

$$c_i^l = \sigma \left(b_i^l + \sum_j c_j^{l-1} \times k_{ij}^l \right) \quad (2)$$

where, c_i^l represents the i^{th} feature in the l^{th} layer; c_j^{l-1} denotes the j^{th} feature in the $(l-1)^{\text{th}}$ layer; k_{ij}^l represents the kernel linked from i^{th} to j^{th} feature, b_i^l denotes the bias for this feature and σ represents the activation function. In the proposed work, two conv1D layers are used. The first convolution layer has 16 convolution kernels and the second convolution layer has 32 kernels, each with size $v = 5$ and shift / stride = 1 in both the layers. The output of conv1D layer (I) is given as input to the conv1D layer (II). The length of the output of convolution layer is generally given by $N - v + 1$ for stride of 1, where N is the corresponding input length. The convolution operation with aforementioned parameters is illustrated in Fig. 4 [41], where,

$$c_1 = k_1x_1 + k_2x_2 + k_3x_3 + k_4x_4 + k_5x_5;$$

$$c_2 = k_1x_2 + k_2x_3 + k_3x_4 + k_4x_5 + k_5x_6;$$

⋮

$$\text{Finally, } c_{N-v+1} = k_1x_{N-4} + k_2x_{N-3} + k_3x_{N-2} + k_4x_{N-1} + k_5x_N.$$

The filter weights, i.e., $k_1 \dots k_5$, are initialized using the He uniform initializer [44] and the bias vector is initialized to all zeros. This operation is performed for each filter in both the layers, hence, there are 16 outputs of conv1D layer (I) and 32 outputs of conv1D layer (II). Since N is taken as 5000 in Fig. 3, the output dimensions of conv1D layers (I) and (II) are 4996×16 and 4992×32 respectively. Let lc denote the length of final output of the convolution layers, which is 4992 here.

- **Activation function (σ):** It plays an important role in capturing the non-linearity of the input signal. Here, rectified linear unit (ReLU) is used as the activation function which is defined as [21]:

$$\sigma(c) = \max(0, c) \quad (3)$$

B. 1D-max pooling layer (Maxpool1D)

The output feature maps (convolution outputs, c) produced from the conv1D layers are given as an input to the 1D max pooling layer, which reduces the feature map dimension by retaining only the maximum value of feature map in a window/

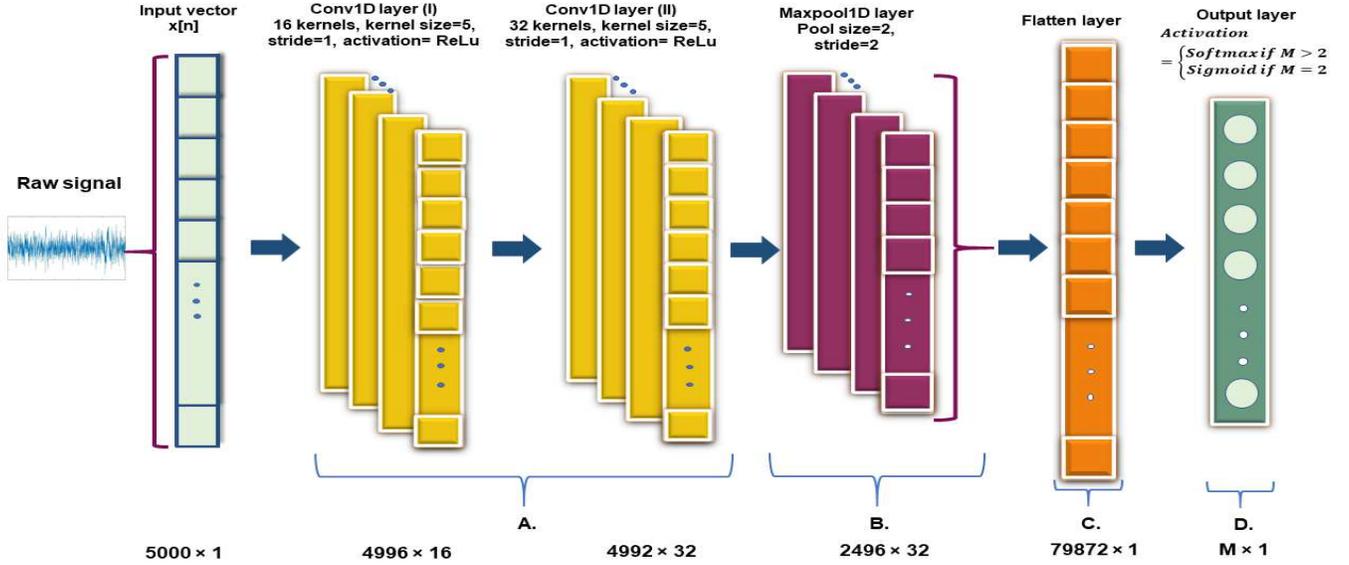


Fig. 3: The proposed CNN architecture for classification of different mental tasks. Note: M denotes the number of classes.

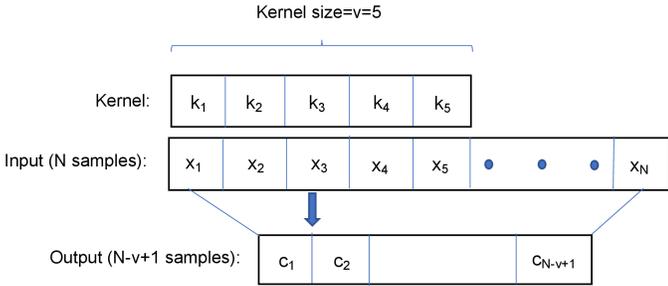


Fig. 4: Illustration of convolution operation at conv1D layer.

patch with a specified pool size [21]. The window is moved across the feature map with a shift/ stride. The operation of max pooling can be represented as [41]:

$$c_h^l = \max_{\forall p \in r_h} c_p^{l-1} \quad (4)$$

where, r_h denotes the pooling region with index h .

In this work, the value of pool size and stride is taken as 2. An illustration of the max pooling operation with these parameters is given in Fig. 5, where, $c_{m_1} = \max(c_1, c_2)$;

$$c_{m_2} = \max(c_3, c_4);$$

$$c_{m_3} = \max(c_5, c_6);$$

⋮

$$c_{m_{l_c/2}} = \max(c_{l_c-1}, c_{l_c}).$$

Hence, the output of this layer has the dimension of 2496×32 which can be seen in Fig. 3.

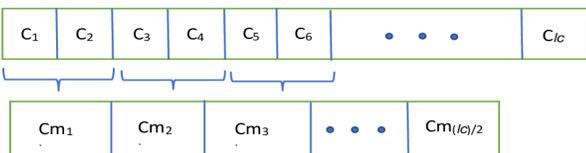


Fig. 5: Illustration of max pooling operation with both pool size and stride of two.

C. Flatten layer and dropout

The flatten layer transforms the input data into a one-dimensional vector, to be fed to the fully connected/ dense layer as shown in the Fig. 3. A dropout parameter is added after the flatten layer, which helps the architecture to generalize well by reducing over fitting during the training process [42]. This is achieved by randomly setting the activations of some nodes to zero, specified by a dropout rate. In this work, a dropout rate of 0.25 has been used.

D. Dense layer for classification

The flattened output is given as an input to the next layer, i.e., dense/fully connected layer which produces the classification output with dimension $M \times 1$, where M denotes the number of classes. In general, the layer operation can be represented as:

$$output = \sigma(\langle input, w_d \rangle + b_d) \quad (5)$$

where, $\langle input, w_d \rangle$ denotes the dot product between weight vector w_d used in this layer and the input, b_d represents the bias vector for this layer and σ is the activation function. In this work, we use both sigmoid and softmax activation for binary and multi-class classification respectively. The sigmoid activation function is given by [43]:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

This function produces the binary output as the probability value for binary classification, based on which the class label is either '0' or '1'. Also, the softmax activation function can be given as [41]:

$$\text{softmax}(\mathbf{z})_i = p_i = \frac{\exp(z_i)}{\sum_{j=1}^M \exp(z_j)} \quad (7)$$

where, z_i represents the i^{th} element of the output vector of previous layer \mathbf{z} . The numerator is normalized by the sum

of all exponential terms from 1 to M to bring the value of p_i between 0 and 1. This layer produces the categorical class labels for multi-class classification. In this work, no bias vector has been used for this layer and the weights are initialized using the glorot uniform initializer [44].

IV. RESULTS AND DISCUSSIONS

In this section, the performance of the proposed architecture is evaluated using different artifact-free and artifactual EEG signals taken from publicly available databases and our recorded database.

A. Performance metrics and training parameters

The performance of the proposed architecture is assessed in terms of following performance metrics [27]:

$$\text{Classification accuracy (Accuracy)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

Here, TP denotes true positives which is the number of cases where the actual positive class is correctly predicted to be positive by the model. TN denotes true negatives which is the number of cases where the actual negative class is correctly predicted to be negative by the model. FP denotes false positives which is the number of cases where the actual negative class is incorrectly predicted to be positive by the model. FN denotes false negatives which is the number of cases where the actual positive class is incorrectly predicted to be negative by the model. For example, in binary classification problem of datasets E and R, we have labeled the non-mental (BT) task as negative and mental (ST) task as positive. If the model predicts a BT task correctly, then it is a true negative. Otherwise, if the model predicts it as an ST task, then it is a false positive. Similarly, if the model predicts an ST task correctly, then it is a true positive, and if it predicts it as BT task, then it is a false negative. Similar interpretation can be drawn for the case of multi-class classification.

$$\text{Precision (PRC)} = TP / (TP + FP) \quad (9)$$

$$\text{Recall (RCL)} = TP / (TP + FN) \quad (10)$$

$$F1 \text{ score} = 2 \times \text{PRC} \times \text{RCL} / (\text{PRC} + \text{RCL}) \quad (11)$$

The performance is evaluated through training and testing of the proposed architecture for the identification and classification of mental tasks. Similar to existing works [4], [24], EEG signal of 10sec duration has been used as an input to the first layer of the model. We perform both binary classification and multi-class classification using the same architecture. In this work, following tasks have been classified: BT-MT, BT-LT, BT-RT, BT-VT, MT-LT, MT-RT, MT-VT, LT-RT, LT-VT, RT-VT, BT-ST and BT-MT-LT-RT-VT from all the three databases. To evaluate the performance of the proposed architecture, 80% of the data is chosen for training and 20% for testing. 20% data for testing is further split into 10% each for testing and validation. Since different subjects have varying number of trials recorded on separate timings/days, 80% of the trials have been selected randomly for training and rest 20% for testing. For training, a batch size of 50, and 20 number of epochs

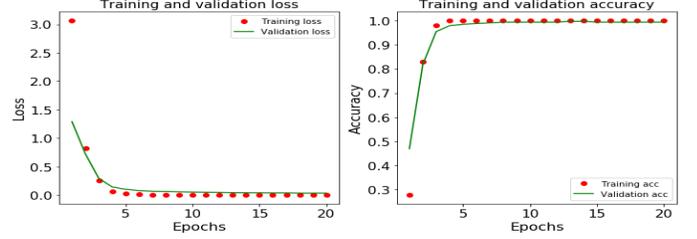


Fig. 6: Training curves for multi-task classification (BT-MT-LT-RT-VT) in database K .

have been used along with the Adam learning algorithm with a learning rate of 0.001. Ten fold cross-validation has been performed for all the three databases. Binary cross entropy and categorical cross entropy are used as loss functions for binary and multi-class classification respectively. These functions are defined as [45]:

$$\text{Binary cross entropy} = -(y \log(p) + (1-y) \log(1-p)) \quad (12)$$

$$\text{Categorical cross entropy} = - \sum_{c=1}^M y_{o,c} \log(p_{o,c}) \quad (13)$$

where, \log represents natural log, y represents binary indicator (0 or 1) if class label c is the correct classification for the observation o , p represents the predicted probability that the observation o is of class label c , M represents the number of classes.

B. Performance analysis

In this section, we demonstrate the classification performance results of the proposed architecture. Figs. 6, 7 depict the training curves with respect to validation and training loss, and validation and training accuracy for the proposed architecture which demonstrate the learning process for the multi-task classification and pair-wise mental task classification in the database K respectively. It can be observed from the curves that the proposed architecture has learnt from the given data in few epochs and does not over-fit. Table II depicts the performance of the proposed architecture in terms of aforementioned performance metrics for all classification tasks and databases. It can be observed from the table that a subject-independent accuracy of 100% has been achieved for the following task pairs: BT-MT, BT-LT, BT-RT, MT-LT, MT-RT, and LT-RT, of the database K . It means that these tasks are accurately classified. It can be observed from the table that the other metrics PRC, RCL are also equal to 1 for these task pairs, which implies that there are no false positives and false negatives respectively. This results in a perfect balance between PRC and RCL, as observed from the F1 score values (which are also 1) for these tasks. Further, an overall average subject-independent accuracy of 99.7% has been achieved for the case of multi-class classification for this database. For databases E and R , proposed architecture achieves an average subject-independent accuracy of 99% and 98% for the classification of BT-ST task pair.

C. Impact of EEG processing length and number of conv1D layers

For assessing the sensitivity performance of the proposed architecture with respect to the processing length of the input

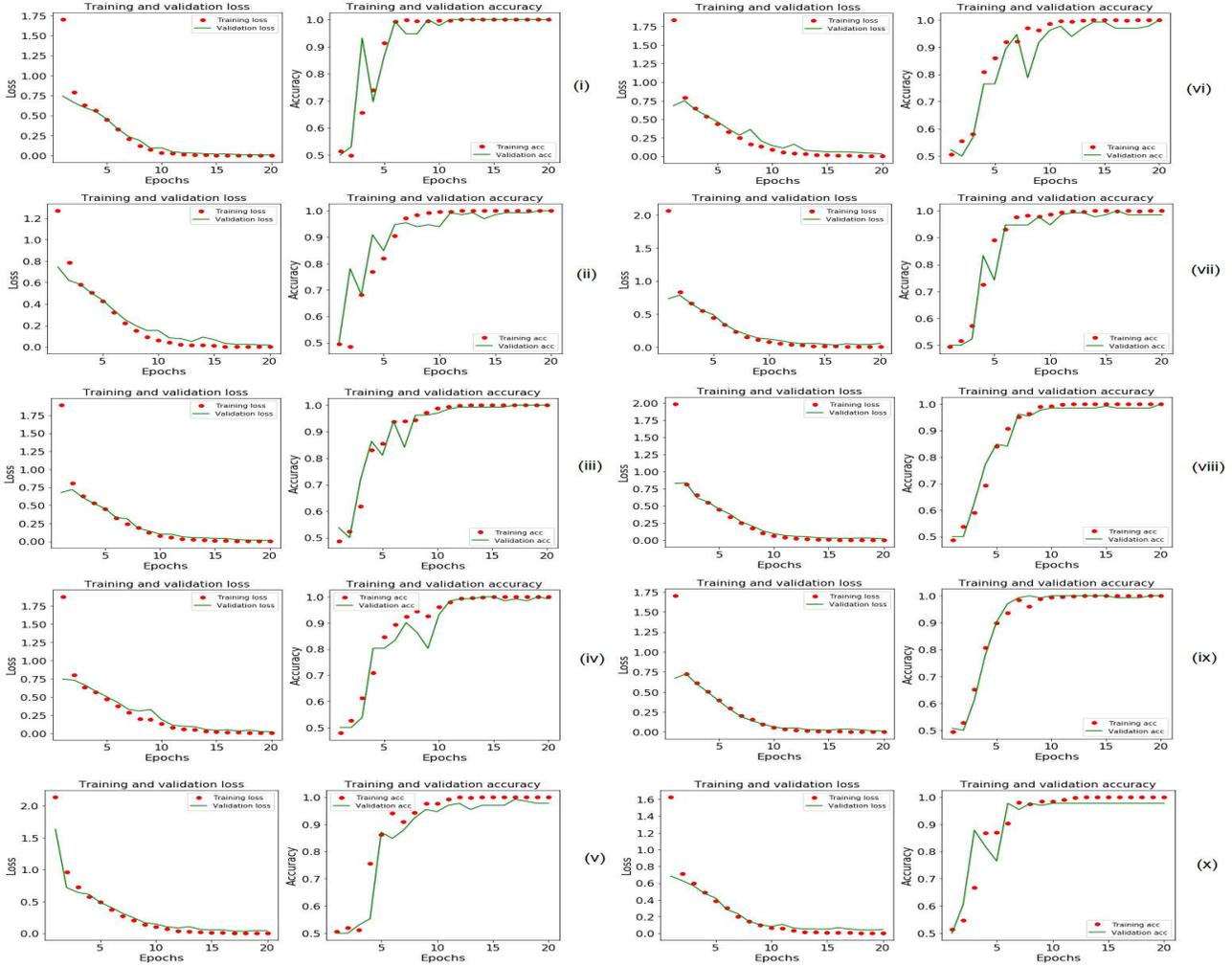


Fig. 7: Training curves for pair-wise mental task classification in database K , (i) BT-MT, (ii) BT-RT, (iii) MT-LT, (iv) MT-VT, (v) LT-VT, (vi) BT-LT, (vii) BT-VT, (viii) MT-RT, (ix) LT-RT and (x) RT-VT.

TABLE II: Mental task classification performance of the proposed architecture for all databases (Mean(std) over subjects (10s data)).

Database	Task	K										E	R	N	
		BT-MT	BT-LT	BT-RT	BT-VT	MT-LT	MT-RT	MT-VT	LT-RT	LT-VT	RT-VT	BT-MT-RT-LT-VT	BT-ST	BT-ST	BT-ST
Accuracy		1 (0)	1 (0)	1 (0)	0.99 (0.12)	1 (0)	1 (0)	0.99 (0.12)	1 (0)	0.98 (0.40)	0.98 (0.21)	0.997 (0.11)	0.99 (0.11)	0.98 (0.22)	0.99 (0.16)
PRC		1 (0)	1 (0)	1 (0)	0.99 (0.20)	1 (0)	1 (0)	0.99 (0.11)	1 (0)	0.98 (0.25)	0.98 (0.20)	0.99 (0.18)	0.99 (0.15)	0.98 (0.27)	0.99 (0.22)
RCL		1 (0)	1 (0)	1 (0)	0.98 (0.11)	1 (0)	1 (0)	0.99 (0.11)	1 (0)	0.98 (0.22)	0.98 (0.11)	0.99 (0.10)	0.99 (0.17)	0.98 (0.21)	0.99 (0.21)
F1		1 (0)	1 (0)	1 (0)	0.98 (0.10)	1 (0)	1 (0)	0.99 (0.11)	1 (0)	0.98 (0.32)	0.98 (0.12)	0.99 (0.16)	0.99 (0.10)	0.98 (0.20)	0.99 (0.17)

signal and the number of conv1D layers, different input lengths ranging from 2s-10s with a step of 2s and varying number of conv1D layers have been used. It is quite observable from the Fig. 8a that the proposed architecture is not significantly variant to the input signal length. Even for shorter durations of processing length, the proposed architecture achieves similar accuracies for different mental tasks classifications. This is useful for the systems where quick response is needed, for example, BCI and neurofeedback systems. Further, it can be observed from the Fig. 8b that the mental tasks classification accuracy is higher for two conv1D layers as compared to one or more than two layers in the proposed architecture. Therefore, two conv1D layers are optimal in the proposed architecture for the classification of mental tasks.

D. Robustness evaluation under different artifacts

In this subsection, we evaluate the robustness of the proposed architecture under different artifacts in the EEG signal. As mentioned earlier, database E contains artifact-free EEG signals and databases K and R consist of EEG signals contaminated with ocular artifacts and muscle artifacts. Hence, in order to evaluate the robustness of the proposed architecture on a large pool of contaminated EEG data, we create a new database namely ' N ' by including only artifact-contaminated EEG signals from database R and all artificially contaminated EEG signals of database E after adding OAs and MAs taken from existing publicly available databases. These OAs and MAs are taken from Mendeley database [46] and MIT BIH polysomnographic database [35], and examples

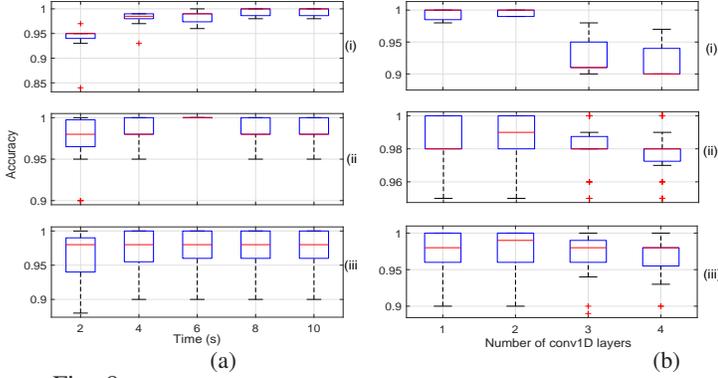


Fig. 8: Illustrates average classification accuracies for (a) varying lengths of input signal and (b) varying number of conv1D layers in: (i) database K , where the accuracies are shown with respect to pair-wise mental task classification, (ii) and (iii) databases E and R with respect to all the subjects respectively.

of electromyogram database [35] and cerebral vasoregulation in elderly with stroke (CVES) database [35] respectively. Various realizations of the contaminated data are generated by randomly adding OAs and MAs separately, as well as both together, to each signal of the database E , after re-sampling all OAs and MAs signals to the rate of 500Hz. Let $x_{cl}[n]$ denote an EEG signal from the database E , $y_o[n]$ denote an OA signal, and $y_m[n]$ denote a MA signal. Then, this process can be summarized as follows: $x_o[n] = x_{cl}[n] + \lambda \cdot (y_o[n])$, $x_m[n] = x_{cl}[n] + \beta \cdot (y_m[n])$, and $x_{om}[n] = x_{cl}[n] + \lambda \cdot (y_o[n]) + \beta \cdot (y_m[n])$. Here, $x_o[n]$ represents OA-contaminated signal, $x_m[n]$ denotes MA-contaminated signal, and $x_{om}[n]$ represents combined OA-MA-contaminated signal. λ is a parameter which denotes the contribution of OAs and β denotes the contribution of MAs. These parameters alter the signal to noise ratio (SNR) of the realized signal [47]. For example, the SNR for OA contaminated signal can be obtained as [47]:
$$\text{SNR}_{(x_o[n])} = \frac{\sqrt{\frac{1}{n} \sum_n x_{cl}^2[n]}}{\sqrt{\frac{1}{n} \sum_n (\lambda \cdot y_o[n])^2}}$$

Similarly, the SNR values for other realizations of contaminated data can be obtained. In this work, the value of λ and β is chosen as 1 such that the EEG signal is fairly contaminated by artifacts. Fig. 9 depicts the training curve with respect to validation and training loss, and validation and training accuracy for the proposed architecture which demonstrate the learning process for classifying BT-ST task in the created database N . In order to demonstrate the robustness of the proposed architecture, mental classification accuracy is computed for contaminated EEG signals with different SNR values ranging 0.4 - 3 which are calculated based on different values of λ and β . Fig. 10 depicts the average classification accuracy at different SNRs which demonstrates the robustness of the proposed architecture in classifying mental and baseline tasks for both artifact-free and artifact-contaminated EEG signals. Also, the average performance of the proposed architecture in terms of performance metrics is demonstrated in Table II for the created database N . As mentioned earlier, the average accuracy is computed for the value of $\lambda = \beta = 1$ and corresponding SNR value of 0.8, as shown in Fig. 10. It can be observed from the table that the proposed architecture achieves an average

accuracy of 99% for the artifact-contaminated database N .

E. Performance comparison

In this subsection, the supremacy of the proposed architecture for mental task classification is analyzed with respect to the other existing mental task classification techniques. Table III demonstrates the overall comparison of the proposed architecture with the existing mental task classification techniques. It can be observed from the table that the proposed architecture not only outperforms existing approaches in terms of accuracy for mental tasks classification but also does not use an artifact suppression/ removal step before training unlike existing approaches. Furthermore, it can be seen from the table that the proposed architecture can accurately classify not only pair-wise mental/baseline task but also five multi-tasks simultaneously unlike existing approaches. To

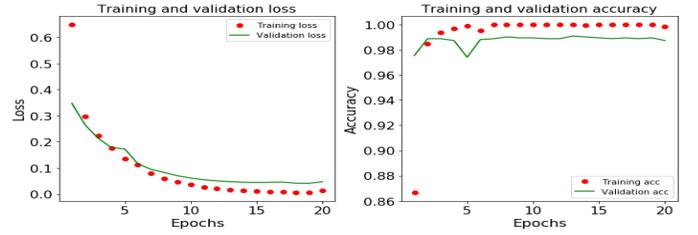


Fig. 9: Training curves for BT-ST task classification in created artifact-contaminated database N .

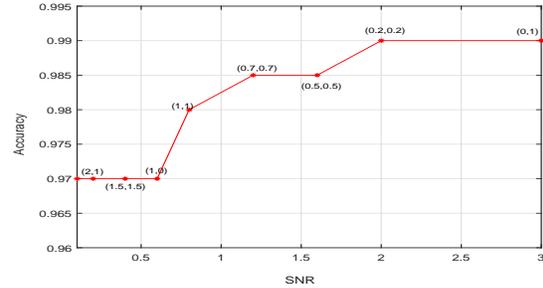


Fig. 10: Illustrates the classification accuracies at different values of (λ, β) and corresponding SNRs for the realized contaminated signals in database N .

demonstrate the learning process through both the layers, the feature/activation maps are produced in the proposed 1D-CNN architecture. The feature maps are extracted as filter weights for conv1D and maxpool1D layers, based on which the outputs of flatten layer and softmax layer correspond to the particular class (mental task). Figs. 11-12 depict the features extracted by the proposed 1D-CNN from EEG signals with baseline and different mental tasks taken from database K and N respectively. From the figures, it can be observed that for each class, the feature maps have inhibitory (small) or excitatory (large) weights for different EEG signals which is illustrated by lighter squares in the Conv1D and Maxpool1D filters. From 11 (a), (b) it can be seen that more filter weights are excitatory in the case of BT, as compared to MT. Also, for the case of RT as shown in Fig.11(d), the activations in maxpool1D are higher for all filters, indicating the brain is in attention state throughout the duration, which is true since the task involves memorizing the figure and visualization of the figure rotation. Similar interpretation can be drawn from

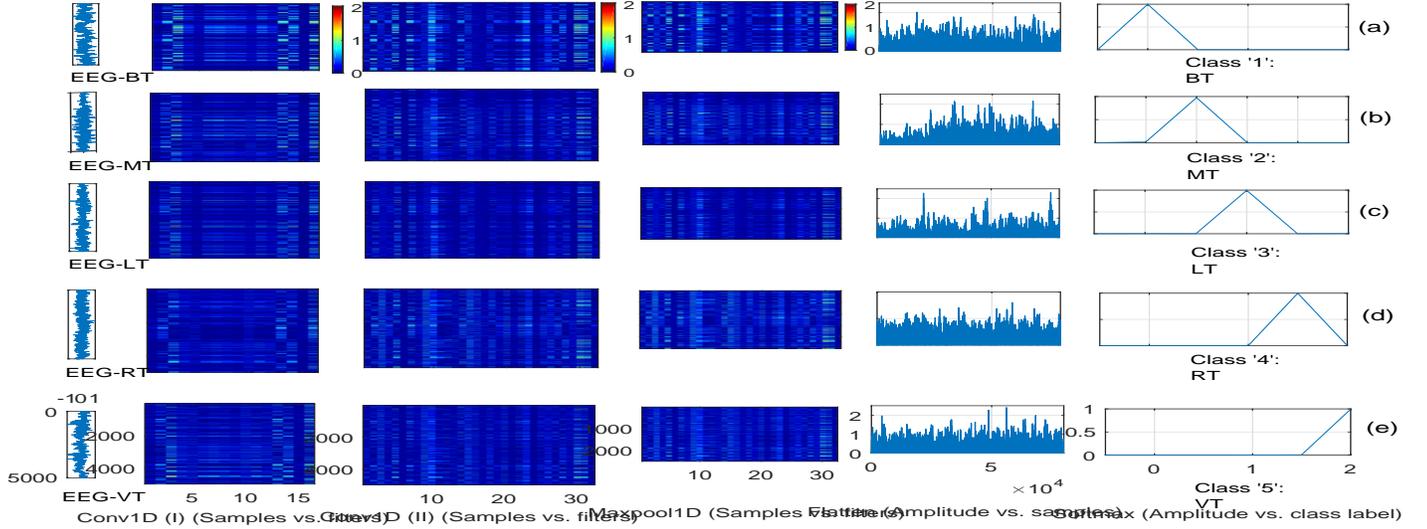


Fig. 11: Feature/activation maps of 1D-CNN layers for signals taken from database K , acquired during different mental tasks.

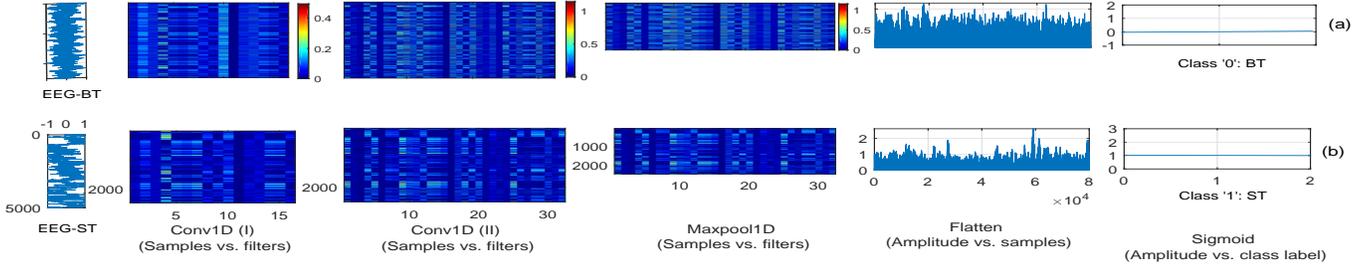


Fig. 12: Feature/activation maps of 1D-CNN layers for signals taken from database N acquired during BT and ST.

TABLE III: Comparison of the proposed architecture with the existing approaches.

Method (Author)	Database	Channel locations (Number)	Subjects (Number)	Artifact removal/rejection	Length used	Accuracy (%)
BQC (Keim and Aunon) [6]	K	$C_3, C_4, P_3, P_4, O_2, O_2$ (6)	5	Y	2s	85-95
FF-NN (Anderson et. al.) [23]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	4	Y	10s	38-71
IF-SVM (Gao et. al.) [27]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	7	Y	10s	88-98
LDA, QDA, SVM (Gupta et. al.) [4]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	6	N	10s	85
KNN, LDA (Noohadi et. al.) [15]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	7	Y	10s	97
EEN-RBP (Pahaniappan) [12]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	4	Y	10s	80-94
LS-SVM (Duna et. al.) [14]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	3	N	10s	94
BP-MLANN (Alyasseri et. al.) [25]	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	7	Y	10s	78-87
SVM-MLP (Anand et. al.) [26]	Self recorded	$F_3, F_2, F_4, C_3, C_2, C_4, P_3, P_4, O_1, O_2$ (9)	41	N	5 min	73
IPSO-NN (Lin et. al.) [24]	IDIAP BCI-III	$C_3, C_4, P_3, P_4, C_{P1}, C_{P2}, P_3, P_4, P_4$ (8)	3	N	4 min	69
SVM-RBF (Wang et. al.) [5]	Self recorded	$AP_3, F_7, F_8, F_{C2}, T_7, P_1, O_1, O_2, P_8, T_8, F_{C2}, F_4, F_8, AP_4$ (14)	10	Y	2 min	97
Random forest (Z. Pei et. al.) [7]	Self recorded	62 EEG channels	7	Y	2s	85
Proposed	K	$C_3, C_4, P_3, P_4, O_1, O_2$ (6)	6	N	10s	99.7 (MT), 99 (TP)
	E	$F_{P1}, F_{P2}, F_3, F_4, F_5, F_7, F_8, C_3, C_4, C_5, P_3, P_4, P_5, O_1, O_2, T_3, T_4, T_5, T_6$ (19)	36	N	10s	99
	R	F_{P1} (1)	20	N	10s	98

BQC: Bayesian quadratic classifier, FF-NN: Feed forward neural network, IF-SVM: Immune feature weighted SVM, QDA: Quadratic discriminant analysis, KNN: K nearest neighbor, BP-MLANN: Backpropagation-multilayer artificial neural network, SVM-RBF: SVM-radial basis function, Y: Yes, N: No.

the feature maps of other mental tasks. Further, it can be observed from the Fig.12 (a), (b) that the activations for BT are higher and uniform for all layers, while for ST, activations are higher in only certain locations. These uniform activations in BT indicate the subject's relaxed mental state as against

random patterns of activations in ST which indicate that the subject has concentrated on some mental activity. Therefore, these features can be mapped to the neurophysiology of the brain. Our proposed CNN architecture has adequately learned discriminatory feature maps for baseline and different mental task classification as depicted by different activation weights which yield high accuracy. To demonstrate the subject adaption of the proposed architecture, we train the proposed model on all signals from EEGMAT and test on the subjects from recorded data. A classification accuracy of 97% has been achieved in this case, which is quite high and indicates the subject-adaptability of proposed architecture. While, the existing features fail in capturing the subject-adaptability, as seen from the performance comparison results. In the future direction, we will analyze the performance of the proposed architecture in case of missing EEG samples and implement the proposed architecture on a real-time embedded processor to determine the real-time latency and power consumption in mental task classification.

V. CONCLUSION

In the proposed work, a light-weight one-dimensional convolutional neural network (1D-CNN) architecture is proposed for mental task identification and classification. The proposed architecture consists of a few layer network which does not require any manual feature extraction or artifact suppression step. The proposed architecture is evaluated using two publicly available databases (i.e, Keim and Aunon (K) database and

EEGMAT (E) database) and in-house database (R) recorded using single-channel neurosky mindwave mobile 2 (MWM2) EEG headset for performing mental/baseline binary classification and mental/mental multi-tasks classification. The proposed architecture achieves classification accuracies of 100% for the following binary task pairs: BT-MT, BT-LT, BT-RT, MT-LT, MT-RT, and LT-RT, in the database K . Further, the proposed architecture achieves an overall average accuracy of 99.7% for multi-class classification in database K , and 99% and 98% for the classification of BT-ST task pair in the databases E and R respectively. Comparative performance results show that the proposed architecture outperforms existing approaches not only in terms of classification accuracy but also in robustness against artifacts. Further, the proposed architecture provides good classification accuracy for shorter processing length of EEG signals which makes it suitable for BCI systems with neurofeedback.

REFERENCES

- [1] D.P. Subha et. al., "EEG signal analysis: a survey," *J Med Syst*, vol. 34, pp. 195–212, 2010.
- [2] R. Palaniappan, "Utilizing gamma band to improve mental task based brain-computer interface design," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no.3, pp. 299-303, 2006.
- [3] M. Jimnez-Guarneros and P. Gomez-Gil, "Custom domain adaptation: a new method for cross-subject, EEG-based cognitive load recognition," in *IEEE Signal Processing Letters*, 2020.
- [4] A. Gupta et al., "On the utility of power spectral techniques with feature selection techniques for effective mental task classification in noninvasive BCI," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1-13, 2019.
- [5] Q. Wang and O. Sourina, "Real-time mental arithmetic task recognition from EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 2, pp. 225-232, 2013.
- [6] Z.A. Keirn and J.I. Aunon, "A new mode of communication between man and his surroundings," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 12, pp. 1209-1214, 1990.
- [7] Z. Pei et. al., "EEG-based multi-class workload identification using feature fusion and selection," in *IEEE Transactions on Instrumentation and Measurement*, 2020.
- [8] P. Arpaia et. al., "A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 8335-8343, Oct. 2020.
- [9] S.R. Butler and A. Glass, "Asymmetries in the electroencephalogram associated with cerebral dominance," *Electroencephalogr. Clin. Neurophysiol.*, vol. 36, pp. 481-491, 1974.
- [10] A. E. Alchalabi et. al., "FOCUS: Detecting ADHD patients by an EEG-based serious game," in *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1512-1520, July 2018.
- [11] B. Wallace et. al., "EEG/ERP: Within episodic assessment framework for cognition," in *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 10, pp. 2525-2534, Oct. 2017.
- [12] F.M. Al-Shargie, et al., "Mental stress quantification using EEG signals," *International Conference for Innovation in Biomedical Engineering and Life Sciences*, Springer, Singapore, 2015.
- [13] Y. Zhang et. al., "Combined feature extraction method for classification of EEG signals," *Neural Computing and Applications*, vol. 28, no. 11, pp. 3153-3161, 2017.
- [14] S. Dutta et. al., "Automated classification of non-motor mental task in electroencephalogram based brain-computer interface using multivariate autoregressive model in the intrinsic mode function domain," *Biomedical Signal Processing and Control*, vol. 43, pp. 174-182, 2018.
- [15] S. Noshadi et. al., "Selection of an efficient feature space for EEG-based mental task discrimination," *Biocybernetics and Biomedical Engineering*, vol. 34, no. 3, pp. 159-168, 2014.
- [16] P. Bashivan, G. M. Bidelman, and M. Yeasin, "Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity," *European Journal of Neuroscience*, vol. 40, no. 12, pp. 3774–3784, 2014.
- [17] M. Saini, Payal, U. Satija, "An effective and robust framework for ocular artifact removal from single-channel EEG signal based on variational mode decomposition", *IEEE Sensor J.*, vol. 20, no. 1, pp.369-376, 2019.
- [18] M. Saini, U. Satija, M.D. Upadhayay, "Effective automated method for detection and suppression of muscle artefacts from single-channel EEG signal," *IET Healthcare Technology Letters*, vol. 7, no. 2, pp. 35-40, 2020.
- [19] A. Qayyum et. al., "Classification of EEG learning and resting states using 1D-convolutional neural network for cognitive load assesment," *2018 IEEE Student Conference on Research and Development (SCoReD)*, Selangor, Malaysia, pp. 1-5, 2018.
- [20] D. Ravi et al., "Deep learning for health informatics," in *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4-21, Jan. 2017.
- [21] Z. Jiao et. al., "Deep convolutional neural networks for mental load classification based on EEG data," *Pattern Recognition*, vol. 76, pp. 582-595, 2018.
- [22] S. Kiranyaz et al., "1D convolutional neural networks and applications: a survey," *arXiv preprint arXiv:1905.03554*, 2019.
- [23] Anderson et. al., "Classification of EEG signals from four subjects during five mental tasks," *Solving Engineering Problems with Neural Networks: Proceedings of the Conference on Engineering Applications in Neural Networks (EANN-96)*, Turkey, 1996.
- [24] C.J. Lin and M.H. Hsieh, "Classification of mental task from EEG data using neural networks based on particle swarm optimization," *Neurocomputing*, vol. 72, no. 4-6, pp. 1121-1130, 2009.
- [25] Z.A.A. Alyasser et. al., "The effects of EEG feature extraction using multi-wavelet decomposition for mental tasks classification," in *Proceedings of the International Conference on Information and Communication Technology*, pp. 139-146, April, 2019.
- [26] R.S. Anand, Gaurav, and V. Kumar, "EEG-metric based mental stress detection," *Network Biology*, vol. 8, no. 1, pp. 25-34, 2018.
- [27] L. Guo et.al., "Classification of mental task from EEG signals using immune feature weighted support vector machines," *IEEE Transactions on Magnetics*, vol. 47, no.5, pp. 866-869, 2010.
- [28] P. Zhang et. al., "Learning spatial-spectral-temporal EEG features with recurrent 3D convolutional neural networks for cross-task mental workload assessment," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 1, pp. 31-42, Jan. 2019.
- [29] A. Craik et. al., "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, no. 3, 2019.
- [30] Y. Yuan et. al., "A multi-view deep learning framework for EEG seizure detection," in *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 83-94, Jan. 2019.
- [31] D. Zhang et. al., "A convolutional recurrent attention model for subject-independent EEG signal analysis," in *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 715-719, May 2019.
- [32] J.X. Chen et. al., "Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks," in *IEEE Access*, vol. 7, pp. 44317-44328, 2019.
- [33] Z. Bai, Y. Ruizhi, and L. Youzhi, "Mental task classification using electroencephalogram signal," *arXiv preprint arXiv:1910.03023*, 2019.
- [34] I. Zyma et. al., "Electroencephalograms during mental arithmetic task performance," *Data*, vol. 4, no. 14, 2019.
- [35] A. Goldberger et. al., "PhysioBank, physioToolkit, and physioNet: components of a new research resource for complex physiologic signals", *Circulation*, vol. 101, no. 23, pp. 215-220, 2003.
- [36] E. W. Nugroho and B. Harnadi, "The method of integrating virtual reality with brainwave sensor for an interactive math's game," *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Chonburi, Thailand, pp. 359-363, 2019.
- [37] U. Satija, B. Ramkumar, M. S. Manikandan, "Real-Time signal quality-aware ECG telemetry system for IoT-based health care monitoring," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 815-823, June 2017.
- [38] U. Satija, B. Ramkumar, M. S. Manikandan, "Automated ECG noise detection and classification system for unsupervised healthcare monitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 722 - 732, May 2017. I.F.-3.45.
- [39] P. Croce et.al., "Deep convolutional neural networks for feature-less automatic classification of independent components in multi-channel electrophysiological brain recordings," in *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 8, pp. 2372-2380, Aug. 2019.
- [40] D. Peng et. al., "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheel set bearings in high-speed trains," in *IEEE Access*, vol. 7, pp. 10278-10293, 2019.
- [41] M. Pérez-Enciso et. al., "A guide on deep learning for complex trait genomic prediction," *Genes*, vol. 10, no. 7, p. 553, 2019.

- [42] H. Wu, and X Gu, "Towards dropout training for convolutional neural networks," *Neural Networks*, vol. 71, pp. 1-10, 2015.
- [43] S. Elfving et. al., "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning," *Neural Networks*, vol. 107, pp. 3-11, 2018.
- [44] B. Hanin and D. Rolnick, "How to start training: The effect of initialization and architecture," *In Advances in Neural Information Processing Systems*, vol. 31, pp. 571-581, 2018.
- [45] I. Goodfellow et. al., "Deep learning," *MIT press*, 2016.
- [46] M.A. Klados and P.D. Bamidis, "A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques", *Data in Brief*, vol. 8, pp. 1004-1006, 2016.
- [47] W.D. Clercq et. al., "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *in IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2583-2587, Nov. 2006.