An Improved Decision Support System for Identification of Abnormal EEG Signals Using a 1D Convolutional Neural Network and Savitzky-Golay Filtering

Unmesh Shukla¹, Geetika Jain Saxena², Senior Member, IEEE, Manish Shailani³, Anil Singh Bafila¹, Amit Pundir², Member, IEEE, Sanjeev Singh¹

¹Institute of Informatics and Communication, University of Delhi South Campus, Delhi, India
²Department of Electronics, Maharaja Agrasen College, University of Delhi, Delhi, India
³School of Computer and Information Sciences, Indira Gandhi National Open University, Delhi, India

Corresponding author: Sanjeev Singh (e-mail: sanjeev@south.du.ac.in).

ABSTRACT Medical experts employ electroencephalography (EEG) for analyzing the electrical activity in the brain to infer disorders. However, the time costs of human experts are very high, and the examination of EEGs by such experts, therefore, accounts for a plethora of medical resources. In this study, an improved one-dimensional CNN-only system of 25 layers has been proposed to identify abnormal and normal adult EEG signals using a single EEG montage without using any explicit feature extraction technique. Most of the previous systems based on deep learning, that have been proposed to solve this problem, use extremely deep architectures containing very large numbers of layers. This study also presents an independent preprocessing module that has been exhaustively evaluated for optimal parameters with the target of adult EEG signal classification. The achieved accuracy of the proposed classifier as a part of the decision support system is 82.24%, which is a substantial improvement of ~3% over the previous best reported classifier of comparable depth. The system also exhibits significantly higher F1-score and sensitivity as well as lower loss. The proposed system is intended to be a part of an expert system for overall brain health evaluation.

INDEX TERMS Electroencephalography, Machine Learning, Decision Support Systems, Convolutional Neural Networks, Data Preprocessing, Savitzky-Golay Filtering, Optimization

I. INTRODUCTION

Electroencephalography (EEG) is the most pivotal diagnostic technique used for the detection of brain-related conditions. It can be used to distinguish between and analyze conditions such as syncope, migraine variants, epileptic seizures, psychogenic non-epileptic seizures, subcortical movement disorders, encephalopathy, and catatonia. Along with being a technique targeted for such brain analyses, one of the primary advantages of EEG is that it is a non-invasive method that registers electrical activity using wet or dry electrodes attached to the scalp [1]. Hence the electrical activity is detectable on the skin surface in the form of electric potential.

Signal acquisition
For standard EEG acquisition, approximately two to a few dozen (wet or dry) electrodes are placed on the subject’s scalp in standard positions. Voltage fluctuations, usually in the order of tens of microvolts (μV), are then measured between selected pairs of electrodes and highly amplified [1]. Thus a typical EEG recording is a set of many irregular simultaneous tracings, indicating changes in voltage between pairs of electrodes. Due to the non-invasive mode of signal acquisition, there is a possibility of a significant amount of noise further causing a reduction in signal quality. As a result, the EEG signal is generally considered impoverished [2]. This nature of the EEG signal adversely affects the capability to differentiate between the various underlying medical conditions associated with the characteristics of the signal. Therefore, it poses a major challenge to the human expert for the diagnosis of the subject. This challenge is one of the primary motivations of the present work.
EEG as suitable data
The biomedical relevance of EEG signals for a human expert is indubitable. However, working on developing an automated system for the classification of EEG signals demands a rationale to justify the use of these signals as ‘suitable’ data. From the data perspective, recording the brain’s electrical activity using an electroencephalogram is one of the primary methods of data acquisition from the human brain. Furthermore, it exhibits excellent temporal resolution, which supports the argument of using the signals as inputs for an intelligent system that can infer abnormalities at the neural level by analyzing the signal in the time domain.

Normal and abnormal adult EEG signals
For adult EEGs, normality and abnormality have specific semantics. They are defined differently for different age and gender groups. The two fundamental tasks carried out by a human expert to analyze an EEG signal of a subject for normality are (1) Background Recognition; and (2) Transients Recognition. The usual background characteristics can be classified as follows.

1. Alpha rhythm: 8-13 Hz
2. Beta activity: 18-25 Hz, 14-16 Hz, 35-40 Hz
3. Theta activity: 5-7 Hz
4. Reactivity: Effects of provocations or specific physiological changes, especially on alpha rhythm
5. Mu rhythm: Central rhythm of the alpha rhythm (8-10 Hz), usually visible in 17% - 19% of adults

Out of the above characteristics, the most relevant characteristic for a human expert for normal EEG identification is the alpha rhythm [3]. It is also known as the Posterior Dominant Rhythm (PDR). It emerges in the occipital region of the brain when the eyes of the subject are closed. When the subject enters the state of drowsiness, then this rhythm fades away. The evaluation of an EEG as abnormal or normal is largely dependent on the analysis of the reactivity observed in the EEG during the emergence of this occipital alpha rhythm. Therefore, the presence, frequency, and distortion of PDR in EEG are considered the primary criteria for an EEG to be classified as normal or abnormal [2].

A. MOTIVATION
Medical interpretation of an abnormal EEG
An abnormal EEG can correspond to various conditions, including, but not limited to, focal slowing or generalized slowing caused by conditions such as epilepsy, delirium or encephalopathy, dementia, and coma [4]. For biomedical interpretation, exact association is assumed between an abnormality in the waveform of an EEG and a corresponding brain condition [5]. These conclusions consider specific signal characteristics such as sharpes, waves, and spikes that are indicative of a known abnormal pattern but are difficult to observe for a novice EEG interpreter.

Problems in manual interpretation
Concluding abnormality in an EEG is not straightforward. For example, in one case, the abnormal nature of an EEG can be inferred on the basis of long and clear periods of wave and spike activity in the signal called epileptiform discharges. However, the mere presence of such activity is not a sufficient indicator to conclude abnormality. This is because the morphologically concluded epileptiform may be a benign variant not necessary linked with epilepsy [3]. A more common issue with EEGs is the amount of external background noise and internal natural noise, which further makes manual interpretation difficult. Another challenge in the interpretation of abnormal EEGs is the coverage of all kinds of abnormalities [4].

Therefore, any decision support system aimed to identify abnormal EEGs must interpret benign variants, eliminate different types of signal noise, and cover variation of signals corresponding to different abnormalities.

B. RELATED WORKS
Many machine learning based methods have been used to identify EEG signals associated with a single abnormality [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] as shown in Table I. Deep convolutional neural networks have been used to solve the problem of automated detection of Parkinson’s disease using EEGs [15] with an accuracy of 88.25%. A range of machine learning methods have also been previously applied for the automated detection of Alzheimer’s disease using EEGs [7] [8] [16]. A previous study [14] has used a deep CNN to develop a system for automated screening of depression using EEGs that has reported an accuracy of 95.49%. Machine learning methods have been used to detect epilepsy using EEGs in some previous works [9] [10] [17]. Another diagnosis that is universally done using EEGs is the detection of seizures. Deep CNNs and complex mathematical features such as wavelet and Fourier features have been used to detect seizures automatically [11] [12]. The deep CNN solution that uses wavelet features to classify EEG signals as epileptic or normal [6] is 88.7% accurate. Three dimensional optimized convolutional neural networks have also been used for epileptic seizure prediction [18]. Pyramidal 1D CNNs have also been used for the detection of epilepsy [13] with an accuracy of ~ 99%.

The use of CNNs for automated identification of EEG signals associated with any abnormality has also been attempted with a test accuracy of 79.34% on the TUH-EEG Abnormal dataset in [19]. The model used in [19] comprises 24 layers. The best CNN-only solution presented in [20] for abnormal EEG identification on the same dataset gives an accuracy of 76.90%. Although there have been very recent developments in the systems for abnormal EEG identification that present accuracies in the range of 85% to 97% [21], the numbers of layers used in these deep learning architectures are
TABLE I
PREVIOUS WORK ON AUTOMATED INTERPRETATION OF EEG AND COMPARISON WITH PRESENT WORK*

| Reference | Problem | Dataset | Method | Accuracy | Specificity | Sensitivity |
|-----------|---------|---------|--------|----------|-------------|-------------|
| [6]       | Epileptic Seizure Detection | [22] | Mixture of Experts | 0.945 | 0.94 | 0.95 |
| [7]       | Alzheimer’s Detection | [23] | SVM | - | 0.94 | 0.85 |
| [8]       | Alzheimer’s Detection | [24] | Fractality Wavelet-Chaos Methodology | 0.993 | 0.97 | 1.00 |
| [9]       | Epilepsy Detection | [22] | ELMAN ANN | 0.960 | 0.99 | 0.92 |
| [10]      | Epilepsy Detection | [22] | Entropies + HOS + Higuchi FD + Hurst + Fuzzy Classifier | 0.997 | - | - |
| [11]      | Seizure Detection | [22] | Dual-tree complex wavelet-Fourier features + NN Classifier | 1.000 | - | - |
| [12]      | Seizure Detection | [22] | 13-layer CNN | 0.887 | 0.90 | 0.95 |
| [13]      | Epilepsy Detection | [22] | Pyramidal-1D-CNN | 0.991 | - | - |
| [14]      | Abnormal EEG Detection | [3] | 1D-CNN-RNN | 0.822 | - | - |
| [15]      | Abnormal EEG Detection | [3] | Manual Collection | 0.954 | - | - |
| [19]      | Parkinson’s Disease Detection | Manual Collection | 13-layer CNN | 0.882 | 0.91 | 0.84 |
| [18]      | Epileptic Seizure Detection | CHB-MIT [25] + NINC [26] + SRM (Manual dataset) | 3D Optimized CNN | 0.980 | - | 0.99 |

Present Work* | Abnormal EEG Detection | [3] | 25-layer CNN | 0.822 | 0.98 | 0.99 |

in the order of hundreds. Therefore, these systems suffer from high computational load even during testing. The system proposed in this study improves the accuracy of the system presented in [19] by ~3% with minimal change in the number of layers. Table I also shows a survey of existing work on similar systems that have been proposed for the identification of EEG signals associated with all abnormalities [19] [20]. The present work uses the same dataset as [3], [19] and [20]. The performance metrics of the proposed system show an improvement over these systems.

II. MATERIAL AND METHODS

A. DATASET
The TUH-EEG Abnormal dataset [3] [27] was used in the current study. This dataset is a subset of the TUH-EEG Corpus [28]. At the time of this study, this corpus [28] was the largest publicly available dataset of EEGs. The distributions of the dataset, with respect to number of patients and number of EEG sessions, are shown in Tables II, III, and IV.

TABLE II
DISTRIBUTION OF NUMBER OF PATIENTS IN THE DATASET

| Description | Normal | Abnormal | Total |
|-------------|--------|----------|-------|
| Test        | 148    | 105      | 253   |
| Train*      | 1237   | 893      | 2130  |
| Total       | 1385   | 998      | 2383  |

Each recorded EEG session was stored in the European Data Format (EDF). Signal acquisition for these EEG recordings was done at a sampling frequency of 250 Hz. The number of channels captured in each recording ranges from 24 to 36.

B. PROPOSED SYSTEM
The proposed system comprises (i) a preprocessing module, and (ii) a 1D convolutional neural network. The modules of the system are described in the following subsections. The complete system is depicted in Fig. 1.

1) PREPROCESSING MODULE
The EEG signals have been preprocessed using a six step process before being fed to the CNN. Fig. 2 depicts the components of the preprocessing module.

TABLE III
DISTRIBUTION OF NUMBER OF PATIENTS IN THE TRAIN AND TEST SETS

| Test | 253 |
| Train* | 2076* |

* 54 patients in the training set were considered both abnormal and normal.

TABLE IV
DISTRIBUTION OF NUMBER OF EEG SESSIONS IN THE TRAIN AND TEST SETS

| Description | Normal | Abnormal | Total |
|-------------|--------|----------|-------|
| Train       | 1371   | 1346     | 2717  |
| Test        | 150    | 126      | 276   |
| Total       | 1521   | 1472     | 2993  |

* 54 patients in the training set were considered both abnormal and normal.
Step 1 - Montage Selection
The raw input signals contain 24-36 channels, out of which the montage of the posterior temporal to occipital (T5-01) channels has been selected for classification. The studies in [19] and [21] have supported the selection of this montage for automated identification of abnormal EEG signals. Moreover, the T5-01 montage has also been recommended as one of the standard longitudinal bipolar montages to be used for clinical purposes [29].

Step 2 - Segmentation
Each EEG recording in the dataset has been segmented for the first 90 seconds. This segment duration has been determined through a comparative analysis carried out by training the 1D-CNN for segments from 5 to 300 seconds by only changing the dimensions of the input layer of the model. Therefore, each instance of the dataset consists of 22500 frames.

Step 3 - Trimming
Initial frames of most EEG signals in the dataset consist of undesirable extrinsic artifacts. Therefore, the first 1250 frames (trimming threshold = 5 seconds) of each segment have been cut. This trimming threshold has been determined using a comparative analysis of model accuracies by regulating this threshold from 0 through 80 seconds (20000 frames).

Step 4 - Min-Max Scaling
Each EEG in the dataset may have different minimum and maximum values from the others. This causes a significant variation in values across the dataset. The EEG data has been normalized to the range [0, 1] to overcome this variation by using min-max scaling. A significant advantage of this scaling is that the scaled sample values are limited to a range that allows comparatively more straightforward computations using smaller numbers during further processing.

Step 5 - Standard Scaling
Suppose the EEG signals, as input data to the CNN, are not standardized. The performance of the model is likely to worsen in that case, as the data is not similar to the standard normally distributed data. Therefore, the EEG signals have been standardized by centering the samples around their mean by subtracting the mean from the sample values and scaling them to unit variance.

Step 6 – Savitzky-Golay Filtering
The Savitzky-Golay (SavGol) filtering technique uses simplified least square procedures to smooth a signal [30]. It uses convolution to fit successive subsets of adjacent sample values using an N-degree polynomial. Smoothing involves using the N-degree polynomial to approximate or smooth W sample values, where W is the length of the filter window. Therefore, the number of convolution coefficients used for smoothing is W. This technique increases the signal-to-noise ratio without significantly distorting the local behaviors of sample values in the signal. A previous study [31] has compared multiple smoothing techniques and supported the usage of this filtering technique to smooth EEG signals. When different models are fit to the smoothed signals and tested for classification, SavGol filtering has shown the best performance.
2) PROPOSED 1D CNN

The proposed convolutional neural network has 25 layers, as depicted in Fig. 3. It includes 1D convolution layers, max-pooling layers, dropout layers, batch normalization layers, a flatten layer, two dense layers, and a final softmax layer for classification. Since valid padding is used in all the convolutional and pooling layers, the right edge inputs that are not covered by the respective filters at these layers are dropped. Other than the softmax activation used for binary classification at the last layer of the model, all the default activations used in the model are rectified linear activation functions in the form of rectified linear units.

The input layer of the model takes the output of the preprocessing module as its input. The first 1D convolutional layer applies eight filters of kernel size equal to 23 on the input signal. An interval of 3 is used as the stride for the application of filters in this layer. Since a large amount of redundant information is present in the 1D feature maps after applying the first convolution layer on the input signals, an initial dropout layer is used to randomly set the input units to zero with a frequency equal to 0.2. The rest of the inputs are scaled up by a factor of 1.25. This ensures that the sum of all inputs remains unchanged. The next layer in the proposed model is a 1D max-pooling layer. A pool size of 2 ensures that the maximum values in 2-sized windows are selected as the representative values of the window. Therefore, the max-pooling operation provides reduction in the size of the feature map by half. The filter window in this layer has a stride of 2. A batch normalization layer is used next to reparametrize the model by normalizing the output using the standard deviation and mean of the input batch. It removes the harmful effects of internal covariate shift [32]. Another 1D convolutional layer applies 64 filters of kernel size equal to 13 on the inputs with a stride of 1. It is followed by a max-pooling layer that halves the size of the feature map. Applying convolution followed by max-pooling in a deep CNN derives the relevant features through convolution layers, intensifies the features through max-pooling and discards the information that makes lesser relevant features for classification. Two 1D convolution layers and a max-pooling layer are used next for further feature extraction with the same rationale of discarding some amount of irrelevant information. Layer-wise activation maps show the presence of redundant data in the 1D feature maps after applying the previous layer. Therefore, another dropout layer with the same parameters is used. Five 1D convolution layers are used next in the network for feature extraction. A dropout layer then randomly sets the input units to zero with a frequency of 0.1, while scaling the rest of the inputs by a factor of 1.11. The next 1D max-pooling layer reduces the size of the feature map by half. Two 1D convolutional layers are then used that apply 32 and 16 filters of kernel size equal to 3 on the inputs with a stride of 1. They are followed by another 1D max-pooling layer. A second batch normalization layer normalizes the output using the standard deviation and mean of the input batch. The inputs are then flattened to prepare for the dense connections in the following layers. Another dropout layer ensures the removal of redundant information from the feature maps. The outputs of this layer are fed to a densely connected layer of 64 units to proceed towards the final classification. In this layer, the output is calculated using (1).

\[
O = A((I, K) + B) 
\]

Here \(I\) is the input, \(K\) is the kernel, \(B\) is the bias vector, \(A\) is the element-wise activation function, and \(O\) is the layer’s output. The subsequent output is then fed into the last densely connected layer of 2 units. Finally, softmax activation is used to classify the input as either 0 or 1, corresponding to the target normal and abnormal classes of EEG signals respectively.

C. EXPERIMENTAL SETUP

This section describes the training, validation, and testing flows that have been used to perform experiments on the TUH-EEG Abnormal dataset. It also comprises the software
and hardware setups that has been used for all the experiments.

1) PROCESS OF TRAINING, VALIDATION, AND TESTING
The standard process for accuracy estimation that has been used for training, validation, and testing sets is depicted in the flowchart in Fig. 4. For many optimizations, this process has been slightly modified. Description of modification in the experimental setup is provided with the description of the corresponding optimization in the Results and Discussion section.

2) SOFTWARE SETUP
In the present investigation, Python and Anaconda have been used for programming and creating virtual environments respectively. Deep learning tasks have been performed using TensorFlow and Keras libraries. Library functions from scikit-learn and scipy have been used for scaling and filtering operations.

3) HARDWARE SETUP
Table V lists the relevant hardware specifications of the machine that has been used to perform all experiments.

| Hardware | Specifications |
|----------|----------------|
| CPU      | Model - Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz Number of CPU cores - 64 |
| GPU      | Model - Tesla V100-PCIE-32GB Number of GPUs - 2 |
| RAM      | Size - 1 TB |

TABLE VI
COMPUTATIONAL TIME SPENT ON EACH PHASE OF MODEL DEVELOPMENT

| Phase                    | Time (in seconds) |
|--------------------------|-------------------|
| Data loading             | 51.56             |
| Preprocessing            | 11.87             |
| Model training           | 204.22            |
| Prediction (single EEG recording) | 0.12 |

III. RESULTS AND DISCUSSION
The computational times for different phases of the system, and the results of classification, parameter and hyperparameter optimizations, are reported in this section. In addition, results of the comparison of accuracies for different preprocessing techniques for EEG signal classification have been discussed. The performance comparisons reported in this study have been made with [19] as it is the best reported model of comparable depth and complexity in existing literature (Table I) that is trained on the same dataset.

Computational time
The computational times required for data loading and preprocessing, model training, and prediction for a single EEG sample have been reported in Table VI.

Optimization of size of the input signal
In two of the previous studies on the identification of abnormal EEGs [3] [19], a 60 seconds duration of the signal, corresponding to 15000 frames on a sampling frequency of 250 Hz, was considered. However, the rationale for selecting only the first 60 seconds was not specified. Therefore, the present study optimized the number of input frames used for classification by thoroughly evaluating the classification accuracies over a range of input signal durations (number of frames). Fig. 5 depicts the change in test accuracies when the size of the input signal was changed from 5 seconds (1250 frames) to 5 minutes (75000 frames). The results in Fig. 5 indicate that the system attains the peak value of the test accuracy when the number of frames as input is 22500.

FIGURE 4. Process of accuracy estimation for training, validation and testing phases
FIGURE 5. Comparison of Test Accuracy vs. Number of Frames as input to the model
Smoothing of EEG or any biomedical signals, in general, is a sensitive process. The incorrect application of smoothing can filter a random amount of significant information required for classification. By traditional signal processing standards, the signal may become smooth, but crucial information essential for abnormal EEG detection may be lost. Therefore, it is critical to identify and compare reliable smoothing techniques for EEG signals. Analysis of different techniques reveals SavGol and median filtering as the most suitable methods [31]. These filtering methods also respect the local behavior of a signal within a window. They provide flexibility in choosing a window size as per the extent of the local behavior of the signal that needs to be preserved. Fig. 8 compares both the techniques for train, validation, and test datasets. Ideally the test accuracy should not be used as a criterion for selecting the hyperparameters of a machine learning model. However, performance metrics on the test data can be utilized for selecting the best smoothing technique during preprocessing. Based on the overall comparative results (Fig. 8) and the flexibility of the filter to accommodate local behavior, the SavGol filtering technique can be concluded as the better technique.

Parameter optimization for the Savitzky-Golay filter

Window size and polynomial order are the two primary parameters of a SavGol filter. They were optimized by evaluating test accuracies over a range of window sizes and polynomial orders. Fig. 9 shows two 3D plots created to visualize the optimization. Fig. 9(a) is a scatter plot while Fig. 9(b) shows a mesh laid on top of the 3D points in the scatter plot. The mesh displays a generic trend of accuracies with changing window sizes and polynomial orders. The figure shows that the optimized values of window size (W) and polynomial order (N) for the highest accuracy over the test set are 9 and 51 respectively. It depicts an initial drop in accuracy when the filter is applied with small values of polynomial order and window size. But when the polynomial order approaches 7, a sudden increase in the accuracy of the proposed system is witnessed. It is evident from the results that increasing the order of the polynomial beyond 9 does not yield better performance.

Convolution coefficients of the filters, calculated as per the optimized values of W (51) and N (9), are shown in Fig. 10. The figure shows a plot of the filter's coefficients applied to the signal during the preprocessing phase.

Inspection of 100 randomly selected EEGs from the original dataset revealed that few frames in the first quarter of the signal length had a visually observable noise. Therefore, the original signal was trimmed by a specific number of frames to obtain better accuracy. This number is referred to as the trimming threshold. Fig. 6 shows the model’s training and test accuracies on increasing the trimming threshold from 0 to 20000 frames. The best accuracies were observed when the trimming threshold is 1250 frames, corresponding to 5 seconds of the initial signal thereby confirming our assumption. The trend of successive decline in the values of the training accuracies (Figure 6(a)) after 1250 frames offers a solid foundation to conclude that the optimal trimming threshold in this study is 1250 frames.

Advantage of Smoothing

Smoothing of an input signal, in general, reduces the time it takes to train the model. The training time is reduced as smoothing removes noise, thereby causing accuracies to converge in fewer epochs. Fig. 7 compares the number of epochs required by the proposed model for convergence of training accuracy with smoothed and unsmoothed signals.

Selection of Savitzky-Golay Filtering

Smoothing of EEG or any biomedical signals, in general, is a sensitive process. The incorrect application of smoothing can filter a random amount of significant information required for classification. By traditional signal processing standards, the signal may become smooth, but crucial information essential for abnormal EEG detection may be
Effect of SavGol filtering

A brute force approach was applied to evaluate the effect of SavGol filtering on classification performance, by comparing the accuracies of the models trained on filtered and unfiltered inputs. For this comparison, the model architecture in [19] was used. Fifty random pairs of subsets were created in the 80%-20% ratio from the original training set to obtain fifty different training and corresponding validation subsets. The best accuracies were then calculated after training and validating the model independently over the 50 pairs of subsets. Finally, the average of the 50 best accuracies was calculated for training, validation and test data. In the present study, this performance metric has been termed as Average Accuracy. Fig. 11 depicts the comparison of average accuracies of the model when the input was filtered and unfiltered. A similar comparison of maximum accuracies was made to verify the contribution of SavGol filtering to create the most accurate version of the model. Again, the maximum accuracy was obtained when the model was independently trained and validated using the 50 unique training and validation subset pairs. Fig. 12 depicts the relevant results. The results indicate that SavGol filtering improves the overall accuracy of the model.

**Hyperparameter Estimation**

The number and types of layers, layer-wise hyperparameters and model configuration were determined by observing the values of validation accuracy for different values of hyperparameters, intuition from activation maps and inspiration from older models [19]. Intuitive selection of the initial values of hyperparameters for the design of the model was also based on the fact that according to human experts, alpha rhythm is the most significant background characteristic for the classification of an EEG signal as abnormal or normal. Therefore, values of filter sizes and strides were chosen such that the frequency of the alpha rhythm was not lost during feature extraction.
The hyperparameters, the number of parameters learnt and the corresponding output shape for each layer are listed in Table VII. A total of 851,746 parameters have been used in the model. The number of trainable parameters out of these parameters is 851,698, while 48 parameters are non-trainable.

### Optimization of Batch Size

Average accuracies were calculated and compared when the model is trained using batch sizes of 128 and 256. Relevant results are depicted in Fig. 13. The results in the figure indicate that a batch size of 256 yields higher average validation and testing accuracies. The final set of hyperparameters used to configure the training phase of the model are listed in Table VIII. The parameters in the optimizer configuration and the type of loss function used were inspired by previous models [19].

### Performance of the Proposed System

The proposed system of EEG classification has been evaluated against [19] as it is the most accurate model of comparable depth and complexity to the best of our knowledge. The average and maximum accuracies of these models were calculated and compared for the train, validation, and test datasets. The results are depicted in Fig. 14 and Fig. 15. Maximum values of precision, sensitivity, specificity, f1-score, classification accuracies, minimum values of losses for the two systems, and the number of layers and parameters in the models have been presented in Table IX. Based on the reported performance metrics in the table, it can be concluded that the proposed system performs better than [19] except for a loss of 0.02 in precision. All other metrics have significantly improved.

Due to the multitude of levels and range of abnormalities a subject may suffer from and the minute nature of differences between them, the research in this area is challenging and inherits limitations of the domain. One such inherent limitation in EEGs is the amount of noise in the signal due to the non-invasive nature of the recording. The proposed system needs to process the signal further to provide more meaningful data inputs to the classification algorithms. The present study needs to be extended to include classification algorithms for specific kinds of abnormalities in EEG signals.
to further provide incisive assistance to medical experts. Like any other medical decision support system, the system inferring the abnormal nature of an EEG also needs to improve till it is 100% accurate. If the normality of an EEG signal is the positive outcome, then a false positive is a far more worrying medical outcome than a false negative. More research is required to decrease the false positive rate and increase the specificity of the system.

### TABLE VII

**LAYER-WISE INFORMATION FOR THE IMPROVED 1D CNN**

| No. | Type       | Hyperparameters | Number of Parameters | Output Shape |
|-----|------------|------------------|----------------------|--------------|
| 0   | Input      | -                | -                    | 21250 × 1    |
| 1   | Conv1D     | Filters = 8      | 192                  | 7076 × 8     |
| 2   | Dropout    | Rate = 0.2       | 0                    | 7076 × 8     |
| 3   | MaxPooling1D | Pool Size = 2  | 0                    | 3538 × 8     |
| 4   | BatchNormalization | Momentum = 0.99 | 32                   | 3538 × 8     |
| 5   | Conv1D     | Filters = 64     | 6720                 | 3526 × 8     |
| 6   | MaxPooling1D | Pool Size = 2  | 0                    | 1763 × 64    |
| 7   | Conv1D     | Filters = 128    | 24704                | 1761 × 128   |
| 8   | Conv1D     | Filters = 32     | 28704                | 1755 × 32    |
| 9   | MaxPooling1D | Pool Size = 2  | 0                    | 877 × 32     |
| 10  | Dropout    | Rate = 0.2       | 0                    | 877 × 32     |
| 11  | Conv1D     | Filters = 256    | 90368                | 867 × 256    |
| 12  | Conv1D     | Filters = 128    | 360576               | 857 × 128    |
| 13  | Conv1D     | Filters = 64     | 41024                | 853 × 64     |
| 14  | Conv1D     | Filters = 64     | 36928                | 845 × 64     |
| 15  | Conv1D     | Filters = 48     | 46128                | 831 × 48     |
| 16  | Dropout    | Rate = 0.1       | 0                    | 831 × 48     |
| 17  | MaxPooling1D | Pool Size = 2  | 0                    | 415 × 48     |
| 18  | Conv1D     | Filters = 32     | 4640                 | 413 × 32     |
| 19  | Conv1D     | Filters = 16     | 1552                 | 411 × 16     |
| 20  | MaxPooling1D | Pool Size = 2  | 0                    | 205 × 16     |
| 21  | BatchNormalization | Momentum = 0.99 | 64                   | 205 × 16     |
| 22  | Flatten    | -                | 0                    | 3280         |
| 23  | Dropout    | Rate = 0.1       | 0                    | 3280         |
| 24  | Dense      | Units = 64       | 209984               | 64           |
| 25  | Dense      | Units = 2        | 130                  | 2            |

### TABLE VIII

**LAYER-WISE INFORMATION FOR THE IMPROVED 1D CNN**

| S.No | Name           | Values |
|------|----------------|--------|
| 1    | Optimizer Configuration | Name = Adam |
|      |                | beta1 = 0.9 |
|      |                | beta2 = 0.999 |
|      |                | Dec = 1e-3 |
|      |                | Learning rate = 0.00001 |
| 2    | Loss Function  | Categorical cross-entropy |
| 3    | Batch Size     | 256    |
| 4    | Epochs         | 250    |

### TABLE IX

**COMPARISON OF KEY METRICS OF THE PROPOSED SYSTEM WITH [19]**

| Metric                          | [19]       | Proposed system |
|---------------------------------|------------|-----------------|
| Max Training Accuracy (%)       | 80.44      | 83.80           |
| Max Validation Accuracy (%)     | 79.22      | 84.37           |
| Max Test Accuracy (%)           | 79.34      | 82.24           |
| Max Precision                   | 0.92       | 0.90            |
| Max Sensitivity                 | 0.75       | 0.99            |
| Max F1-score                    | 0.74       | 0.80            |
| Max Specificity                 | 0.97       | 0.98            |
| Min Training Loss               | 0.44       | 0.37            |
| Min Validation Loss             | 0.47       | 0.40            |
| Min Test Loss                   | 0.48       | 0.44            |
| Number of Parameters            | 382,498    | 851,746         |
| Number of Layers                | 24         | 25              |

### IV. CONCLUSION

The development of classification models for determining the medical state of a human subject using physiological signals is a thriving domain in current research. Minor improvements in the accuracies of such models are relevant as they help improve the quality of life. However, most previous studies on solving similar problems for physiological signals such as EEGs do not emphasize the preprocessing of signals. The present work underlines the significance of the preprocessing phase by suggesting an independent module for it. Furthermore, we have systematically selected and optimized the components of the preprocessing module by illustrating the stepwise improvement in the classification accuracies. As a result, the maximum test accuracy obtained by the proposed system is 82.24%, a significant improvement (~3%) over the previously reported best test accuracy of 79.34% in the base work [19]. Significant improvements in the average accuracies show that the system is reliably accurate. Higher
specificity, sensitivity and f1-score indicate that the system is more effective for biomedical application than the system proposed in [19].

The proposed decision support system can assist human experts in the initial screening and post-diagnosis validation of EEG recordings. The system is being improved to eventually utilize it as an effective medical expert system that carries out brain health evaluation. In addition, a user-friendly clinical interface for the decision support system is being designed. A hybrid setup amalgamating the proposed system with explicitly derived features is also intended to be built to improve classification accuracy.

V. FUNDING

No funding was received by the authors for this work.

VI. DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

VII. ACKNOWLEDGEMENTS

The authors would like to thank the team of Project Samarth, an initiative of the Ministry of Education, India, at the University of Delhi South Campus (UDSC), for their technical support.

REFERENCES

[1]. A. B. Benevides et al., “Review of the Human Brain and EEG Signals,” in Introduction to Non-Invasive EEG-Based Brain-Computer Interfaces for Assistive Technologies, 1st ed. Boca Raton, FL, United States: CRC Press, 2021, ch. 1, pp. 1–45, DOI:10.1201/9781003049159.

[2]. J. S. Ebersole et al., “Current practice of clinical electroencephalography,” Wolters Kluwer Health, 4th ed. Philadelphia, PA, United States: LWW, 2014. Available: https://www.scribd.com/document/565979581/Current-Practice-of-Clinical-Electroencephalography-4e

[3]. S. López et al., “Automated Identification of Abnormal Adult EEGs,” in 2015 IEEE Signal Processing in Medicine and Biology Symposium, 2015, pp. 1-5. DOI:10.1109/SPMB.2015.7405423.

[4]. J. W. Britton et al., “The Abnormal EEG,” in Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants, Chicago, IL, United States: AES; 2016. Available: https://www.ncbi.nlm.nih.gov/books/NBK390357/

[5]. P. D. Emmady et al., “EEG Abnormal Waveforms,” in StatPearls, Treasure Island, FL, United States: StatPearls Publishing, 2021. Available: https://www.ncbi.nlm.nih.gov/books/NBK557655/

[6]. A. Subasi, “EEG signal classification using wavelet feature extraction and a mixture of expert model,” Expert Systems with Applications, vol. 32, no. 4, pp. 1084-1093, 2007. DOI:10.1016/j.eswa.2006.02.005.

[7]. C. Lehmann et al., “Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG),” Journal of Neuroscience Methods, vol. 161, no. 2, pp. 342–350, 2007. DOI:10.1016/j.jneumeth.2006.10.023.

[8]. M. Ahmadlou et al., “Fractality and a wavelet-chaos-methodology for EEG-based diagnosis of Alzheimer disease,” Alzheimer Disease and Associated Disorders, vol. 25, no. 1, pp. 85–92, 2011. DOI:10.1097/WAD.0b013e3181ed1160.

[9]. H. Işık et al., “Diagnosis of epilepsy from electroencephalography signals using multilayer perceptron and Elman Artificial Neural Networks and Wavelet Transform,” Journal of Medical Systems, vol. 36, no. 1, pp. 1–13, 2012. DOI:10.1007/s10916-010-9440-0.

[10]. U. R. Acharya et al., “Automated EEG analysis of epilepsy: A review,” Knowledge-Based Systems, vol. 45, pp. 147-165, 2013. DOI:10.1016/j.knosys.2013.02.014.

[11]. G. Chen, “Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features,” Expert Systems with Applications: An International Journal, vol. 41, pp. 2391-2394, 2014. DOI:10.1016/j.eswa.2013.09.037.

[12]. U. R. Acharya et al., “Deep convolutional neural network for the automated detection and diagnosis of seizures using EEG signals,” Computers in Biology and Medicine, vol. 100, pp. 270-278, 2018. DOI:10.1016/j.compbiomed.2017.09.017.

[13]. I. Ullah et al., “An automated system for epilepsy detection using EEG brain signals based on deep learning approach,” Expert Systems with Applications, vol. 107, pp. 61-71, 2018. DOI:10.1016/j.eswa.2018.04.021.

[14]. U. R. Acharya et al., “Automated EEG-based screening of depression using deep convolutional neural network,” Computer Methods and Programs in Biomedicine, vol. 161, pp. 103-113, 2018. DOI:10.1016/j.cmpb.2018.04.012.

[15]. S. L. Oh et al., “A deep learning approach for Parkinson’s disease diagnosis from EEG signals,” Neural Computing & Applications, vol. 32, no. 15, pp. 10927–10933, 2020. DOI:10.1007/s00521-018-3689-5.

[16]. N. Kulkarni, “Electroencephalogram and Its Use in Clinical Neuroscience,” in EEG-based diagnosis of Alzheimer’s disease: a review and novel approaches for feature extraction and classification techniques, San Diego, CA, United States: Academic Press, 2018, ch. 1, pp. 1-24. DOI:10.1016/C2017-0-00543-8.

[17]. S. J. Smith, “EEG in the diagnosis, classification, and management of patients with epilepsy,” Journal of Neurology, Neurosurgery, and Psychiatry, vol. 76, no. 2, pp. ii2-ii7, 2005. DOI:10.1136/jnnp.2005.069245.

[18]. B. P. Prathaban et al., “Dynamic learning framework for epileptic seizure prediction using sparsity based EEG Reconstruction with Optimized CNN classifier,” Expert Systems with Applications, vol. 170, pp. 114533, 2021. DOI:10.1016/j.eswa.2020.114533.

[19]. Ö. Yıldırım et al., “A deep convolutional neural network model for automated identification of abnormal EEG signals,” Neural Computing & Applications, vol. 32, no. 20, pp. 15857–15868, 2020. DOI:10.1007/s00521-018-3889-z.

[20]. S. Roy et al., “Deep Learning Enabled Automatic Abnormal EEG Identification,” in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2756-2759, 2018. DOI:10.1109/EMBC.2018.8512756.

[21]. R. Bajpai et al., “Automated EEG pathology detection based on different convolutional neural network models: Deep learning approach,” Computers in Biology and Medicine, vol. 133, p. 104434, 2021. DOI:10.1016/j.compbiomed.2021.104434.

[22]. R. G. Andrezejak et al., “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” Physical Review E, vol. 64, p. 061907, 2001. DOI:10.1103/PhysRevE.64.061907.

[23]. T. Koenig et al., “Decreased EEG synchronization in Alzheimer’s disease and mild cognitive impairment,” Neurobiology of Aging, vol. 26, no. 2, pp. 165–171, 2005. DOI:10.1016/j.neurobiaging.2004.03.008.

[24]. C. Goh et al., “P04.3 Characterisation of EEG at different stages of Alzheimer’s disease (AD),” Clinical Neurophysiology, vol. 117, pp. 138-139, 2006. DOI:10.1016/j.clinph.2006.06.220.

[25]. S. Ali, “Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment,” Ph.D. dissertation, DSspace@MIT, MIT, Cambridge, MA, 2009. Available: https://dspace.mit.edu/bitstream/handle/1721.1/54669/607331139-MIT.pdf
[26] L. Stevenson et al., “A dataset of neonatal EEG recordings with seizures annotations,” Scientific Data, vol. 6, no. 1, p. 190039, 2019. DOI:10.1038/sdata.2019.39.

[27] A. Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,” Circulation [Online], vol. 101, no. 23, pp. E215–E220, 2000.

[28] A. Harati et al., “The TUH EEG CORPUS: A big data resource for automated EEG interpretation,” IEEE Signal Processing in Medicine and Biology Symposium, pp. 1-5, 2014. DOI:10.1097/WNP.0000000000000317.

[29] J. N. Acharya et al., “American Clinical Neurophysiology Society Guideline 3: A Proposal for Standard Montages to be Used in Clinical EEG,” Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society, vol. 33, no. 4, pp. 312–316, 2016. DOI:10.1097/WNP.0000000000000371.

[30] A. Savitzky et al., “Smoothing and Differentiation of Data by Simplified Least Squares Procedures,” Analytical Chemistry, vol. 36, pp. 1627-1639, 1964. DOI:10.1021/ac60214a047.

[31] A. Kawala-Sterniuk et al., “Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes,” Sensors, vol. 20, no. 3, p. 807, 2020. DOI:10.3390/s20030807.

[32] S. Ioffe et al., “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015, vol. 37, pp. 448-456. Available: http://proceedings.mlr.press/v37/ioffe15.pdf

MANISH SHAILANI was born in New Delhi, India, in 1989. He received his B.Tech. and M.Tech. degrees in Information Technology from Guru Gobind Singh Indraprastha University, Delhi, in 2011 and 2013, respectively. He has obtained his Ph.D. degree from Institute of Informatics and Communication, University of Delhi South Campus, New Delhi.

He is currently serving as an Assistant Professor at School of Computer and Information Sciences in Indira Gandhi National Open University, Delhi, India. His research interests include information security, cyber-physical systems and machine learning.

ANIL SINGH BAFILA received his Bachelor of Science degree in Electronics Honors from Delhi, India, in 2012 and M.S. degree in Informatics from University of Delhi, New Delhi, India, in 2014. He is currently pursuing the Ph.D. degree in Informatics at Institute of Informatics and Communication, University of Delhi, New Delhi, India.

He is a Research Assistant with Institute of Informatics and Communication, University of Delhi, New Delhi, India. His current research interests include machine learning, cloud computing and Internet of Things.

AMIT PUNDIR received his M.Phil. degree in Electronics from the University of Delhi in 1993 and a Ph.D. degree in Electronics from the University of Delhi South Campus in 1998. His doctorate was in the field of semiconductor materials, fabrication and characterization. He is currently a faculty member in Department of Electronics, Maharaja Agrasen College, University of Delhi and Member, IEEE.

He was in the Institute of Life Long Learning (ILL), University of Delhi as a content writer for electronics from 2007 to 2008. His research interests include data analytics and visualization, machine learning, deep learning and cyber-physical systems.

MANISH SHAILANI is a Research Scholar at the Institute of Informatics and Communication, University of Delhi South Campus, India. He received the degrees of B.E. in Computer Science and Engineering from Jabalpur Engineering College, India in 2014 and M.Tech. in Information Technology with specialization in Human Computer Interaction from Indian Institute of Information Technology Allahabad, in India in 2016.

From 2016 to 2019, he has worked in the domain of delivering healthcare services such as EMRs, Diagnostics, Medicine Orders and E-consultations using IT. Presently, his research is focused on biomedical applications of machine learning. His current research interests are physiological signal processing, biomedical informatics, emotion recognition and medical expert systems.

GEETIKA JAIN SAXENA is an Associate Professor at Department of Electronics in Maharaja Agrasen College, University of Delhi, India and Senior Member, IEEE. She received her B.Sc. and M.Sc. degrees in Electronics in 1998 and 2000, respectively. She did her doctorate in Modeling and Simulation of Optical Amplifiers in 2010 from University of Delhi, India.

Her current research interests are digital image processing, machine learning applications and health informatics. She is also working on design, modeling and simulation of optical components for optical communication systems, analysis and design of integrated optical waveguide and devices. She has around 40 publications in international and national peer reviewed journals and conferences to her credit.

UNMESH SHUKLA is a Research Scholar at the Institute of Informatics and Communication, University of Delhi South Campus, India. He received the degrees of B.E. in Computer Science and Engineering from Jabalpur Engineering College, India in 2014 and M.Tech. in Information Technology with specialization in Human Computer Interaction from Indian Institute of Information Technology Allahabad, in India in 2016.

SANJEEV SINGH has worked in optical electronics and obtained his Doctorate in Electronics from the University of Delhi in 1997 from the University of Delhi. Apart from teaching M.Sc. in Informatics for the last twenty years, he has been working in networks, security and communication technologies. He has published over 70 research publications in international/national journals and conferences. He has been a visiting fellow at the University of Manchester in Science and Technology (UMIST), the UK, under Commonwealth Fellowship. He spearheaded the project “Cross-Cultural Innovation Network,” sponsored by the European Commission and completed several research projects funded by various government agencies of India.

He is a member of the National IT committee, FICCI and the principal coordinator for the UK India Education and Research Initiative (UKIERI-I) project. Prof. Singh is a Network Fellow of Commonwealth Partnership in Technology Management (CPTM), London, UK and a member of the Innovation Cluster and Science & Technology Parks (STPs) forum, Ministry of Knowledge Economy, Korea. In addition, he is chair of the national

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
committee on wearable devices at BIS, Govt. of India. His current research interests are IoT-based system design and cyber physical systems.