IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 August 2023, accepted 5 September 2023, date of publication 25 September 2023, date of current version 5 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3319248

RESEARCH ARTICLE

Classifying Parkinson's Disease Using Resting State Electroencephalogram Signals and U^{EN}-PDNet

SYED QASIM AFSER RIZVI[1], GUOJUN WANG[1], (Member, IEEE), ASIF KHAN[2], (Member, IEEE), MOHAMMAD KAMRUL HASAN[3], (Senior Member, IEEE), TAHER M. GHAZAL[3,5,6], (Senior Member, IEEE), AND ATTA UR REHMAN KHAN[4], (Senior Member, IEEE)

[1]School of Computer Science, Guangzhou University, Guangzhou, Guangdong 510006, China
[2]Department of Computer Application, Integral University, Lucknow, Uttar Pradesh 226026, India
[3]Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[4]College of Engineering and IT, Ajman University, Ajman, United Arab Emirates
[5]Applied Science Research Center, Applied Science Private University, Amman 11937, Jordan
[6]School of Information Technology, Skyline University College, Sharjah, United Arab Emirates

Corresponding author: Guojun Wang (csgjwang@gzhu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62372121, and in part by the Key Research and Development Program of China under Grant 2020YFB1005804.

ABSTRACT Parkinson's Disease (PD) in the set of neuro-degenerative disorders stimulates due to the loss of dopaminergic neurons from the substantia nigra. Electroencephalogram (EEG) signals are being extensively utilized for diagnosing PD. The existing approaches extract the features using various frequency transformations that lose valuable signal information. An optimized Deep Convolutional Neural Network (CNN) inspired by the encoder part of U-Net architecture is proposed for classifying PD incorporating the resting electroencephalogram (EEG) signal dataset. The proposed model follows the U-Net architecture for extracting the features from the signals. The EEG recordings are taken from two datasets: the University of Mexico (UNM) EEGs and the University of California San Diego (UCSD) resting state dataset. The EEGs are pre-processed with a basic pre-processing pipeline, then separated into single channels, plotting each channel as a simple graph. These graphs are then fed to the proposed 23-layered convolutional neural network (CNN) for classifying PD from the normal control. Consequently, the model achieved maximum values of 93.10%, 93.18%, 93.09%, and 0.9313 of accuracy, precision, recall, and F1-score respectively for the UNM dataset, whereas, 97.90%, 98%, 97.87% and 0.9794 of accuracy, precision, recall, and F1-score respectively for UCSD dataset. The results show improved scores compared to the individual Machine Learning and CNN models applied on the same datasets.

INDEX TERMS CNN, EEG, PD, deep learning.

I. INTRODUCTION

Parkinson's Disease (PD) is a well-known neurodegenerative disorder influencing a substantial population [1]. It progresses in a manner impacting the motor actions of the patient intensely by inducing the loss of dopaminergic neurons in the substantia nigra pars compacta [1], [2]. It deprives the patient of cardinal motor activities, including cognitive and speech control too [3]. Some early signs tell about the beginning of neurodegeneration; however, PD is generally signified by two broad classes of symptoms, wittingly, motor and non-motor symptoms [3], [4]. Non-motor symptoms include a lot of indications, but the major contributor towards the diagnosis or prognosis of PD is the Motor Symptoms in contrast with neuroimaging techniques [5]. Neuroimaging translates cognitive functionality extensively, analyzing the

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

varied transformations occurring in the brain utilizing images and signals [6], [7]. Whether they are chemical reactions or electrical impulses, explaining the motor actions precisely [3], [5].

Quantitative neuroimaging techniques are currently in progress to decipher the brain. In accordance, a lot of functionality has been traced till now, including the malfunctioning areas damaging motor activities of the body. Significantly, motor actions, particularly are the best source to explain the reasons for the deficiencies happening in Central Nervous System (CNS). Diagnosing PD too relies on the motor symptoms that include Tremors, Rigidity, Bradykinesia, Postural Instability, Difficulty in Walking (Gait), Dystonia, and Vocal System [1], [3], [8]. Neuroimaging techniques demonstrate motor functionalities to a greater extent, including Computed Tomography (CT) scans, Magnetic resonance imaging (MRI), Functional magnetic resonance imaging (fMRI), Brain-Computer Interface (BCI), etc., and the signals like Electroencephalography (EEG) [9], [10]. The EEG signals to read the electrical impulses produced by the chemical reaction within the brain, describing the specific action proposed by the brain based on certain stimuli. If someone touches any hot appliance immediately, the brain recognizes the instrument is hot and retracts the hand [8], [11].

EEG signals are vastly used brain signals to read the brain's electrical activity. They are used in most neurodegenerative diseases like Alzheimer, Parkinson's, and so on [8]. EEG signals are classified into five basic frequency spectra inclusive of delta (δ) ranging from 0.5 —4 Hz, theta (θ) ranging from 4 —7 Hz, alpha (α) 8 —12 Hz, beta (β) ranging from 14—26 Hz and gamma (γ) above 30 Hz or some authors restricts to 45 Hz [12], [13]. These waves depict the body's functionality, from which the beta wave reflects the active waking state of the body, indicating active thinking, focusing on the worldly creatures and problem-solving state. Analyzing these signal reflexes paves the way for detecting the difficulty in executing particular motor actions. As is though, they are used for detecting some major neurological disorders, including Epileptic disorders, Schizophrenia, Alzheimer and PD [4], [8], [14]. To translate these signals appropriately, different techniques are used. Most of the authors have used frequency transformation to make the signal readable. For classification, machine learning approaches were deployed excessively [8].

Medical data in the form of signals and images is accumulating exponentially; the modalities themselves are making it enormous since the recording is taking GBs to TBs and counting. Processing exhaustive data and making predictions from them is not easy. A triad of computer science known as Artificial Intelligence (AI) is sculpted to aid these types of datasets to map human intelligence with the machines [15]. AI, in turn, contains Machine Learning (ML) that uses machines to analyze the data for making predictions to train the machines [8]. Inspiring the researchers to employ ML, most of the research has been conducted employing state-of-art ML techniques, specifically in the medical domain. Lots of ML techniques have been applied in diagnosing neurodegenerative diseases. Likewise, much of the research is done in diagnosing Epileptic disorders, Alzheimer, and schizophrenia along with depression-related disorders [8], [16], [17], [18]. Correspondingly, extensive efforts are laid to diagnose PD by employing ML over different data sets. Inclusive of hand-written digits, speech dataset, performing different tasks by applying head mount EEG cap [3].

A. MOTIVATION OF STUDY

A lot of research has been concluded regarding neurodegenerative disease. Also, compelling efforts have been laid to diagnose PD using different methodologies, with every mechanism adopting different modalities inclusive of voice dataset, assigning a task to the patient, and gait movements. Each modality is quite an effective way to pave the way toward learning the patterns of PD. To differentiate PD from normal controls, motor activities should be understood precisely. For such consideration, direct cognitive translation is a better way, and EEG is one such technique.

B. MAIN CONTRIBUTIONS

The proposed methodology aims to classify PD patients by employing deep learning focussing on:

- Dealing with the raw EEGs, only with basic pre-processing. Compared to the deep learning method, classical machine learning methods, such as kNN, SVM, RF, etc., require hand-crafted feature engineering for classification. Subsequently, the results may be biased.
  The proposed methodology employs the raw EEG for classifying Parkinson's patients from standard control. Compared to classical machine learning algorithms such as kNN, SVM, and RF, classification requires hand-crafted feature engineering. We have proposed a customized CNN architecture.
  For preprocessing the EEGs we have followed the basic preprocessing steps discussed in the section III-B. These steps are utilized to clean the EEGs and make them more interpretable. The pre-processing steps include noise removal, artifacts removal, filtering, bad channel removal, and additional steps explained in the sections above specific to the resting state EEGs. These steps are followed to make the EEGs suitable for input to the CNN.
  One of the significance of the proposed approach is that we do not have to follow the complex step of manual feature extraction since we have plotted the pre-processed signals in the form of images that could be directly fed to the CNN, which directly learns and can draw the relevant spatial patterns and temporal dependencies for the EEGs.

By transforming EEGs into signals, we have neutralized the limitations and biases bounded with the feature engineering or manually extracting the features, which allows the CNN to leverage the inherent information contained within the raw EEGs.

➢ Applying CNN in its simplest form. Inspired by the encoder part of the U-Net architecture, we have doubled the number of filters in every convolutional layer starting from 32 filters to 256 filters in the fourth convolutional layer.

As we know, the U-Net architecture is well-known for image segmentation. One of the distinctive features of the U-Net architecture is its feature learning technique which empowers the U-Net architecture to segment the image efficiently. The U-Net architecture comprises two main parts, encoder and decoder; the encoder part extracts the essential feature of the image or low-level feature, and in the decoder part, these fine-grained features are combined with some high-level features for understanding the particular area of the image.

We only follow this feature extraction approach to the encoder part since we classify the images rather than the segment. To learn the fine features of the signal, we followed the encoder part of the U-Net architecture. The EEG signals plotted on one figure make it more complex to read; hence we have doubled the number of filters in every convolutional layer to make CNN understand from the low-level features to the high-level features.

In the proposed methodology, we have started with a relatively low number of filters, specifically 32 in the first convolutional layer. Onward, we have increased the number of filters up to 256 in the last means fourth convolutional layer, this increasing order of filters lets the network learn more abstract or complex features as it travels deeper in the network. The sequential increase in the number of filters will help the network learn the input signal's hierarchical impression by acquiring both low-level and high-level features related to the classification task.

The inheritance of the encoder part of the U-Net architecture enhances the proposed custom CNN in its discriminative capability. The successive increase in the number of filters allows the CNN to learn the intricate patterns and the structures existing in the plotted EEG signals improving the classification.

Conclusively, we have inherited the encoder of U-Net architecture into the simple CNN architecture to minimize the complexity of the model while maintaining the balance between the complexity and the performance, facilitating the efficient feature extraction and classification of the EEG signals with its complex structure.

➢ Using limited resources constituting valuable contribution in classifying PD. We have performed all the pre-processing along with training the model on an HP OMEN series laptop having the configuration, ''Intel(R) Core(TM) i5-7300HQ CPU 2.50 GHz processor containing one NVIDIA GeForce GTX 1050 GPU of 2GB''. However, the model consists of 23 layers, with every training data containing more than 1000 readings. The model over-performed state-of-art approaches using the same dataset.

With limited resources on board, the proposed 23-layered architecture achieved promising results. The complexity of the dataset could be considered as one of the two datasets having 60 channels plotted on one figure, with each sub-category having more than 1000 readings. The proposed architecture has learned the complex features with intricate patterns from the EEG signals by successfully classifying the patients from the normal controls.

Our approach demonstrated that training the model does not necessarily require high computing devices; it could be done on limited resources. This perspective is exceptionally considerable since it makes the proposed methodology more convenient and functional in a real-world environment having limited resources.

In addition, the proposed model has shown better performance than the other models operated on the same dataset. Dominant to other models, our approach demonstrated better functionality in the evaluation matrices inclusive of accuracy, precision, recall, and F1-score. Prominent to other discussed approaches operational on the same dataset, the proposed methodology contributes to Parkinson's classification and presents a compelling alternative for researchers in the field.

The rest of the paper is framed in the following manner. Section II discusses some of the latest research in characterizing Parkinson's disease (PD). The proposed framework describing the data acquisition up to the proposed CNN architecture in distinguishing the six cases that emerge from the two datasets is discussed in Section III. The formulated outcome from the proposed architecture is presented in Section IV. Consequently, in Section V, the interpretations made by the proposed methodology, inclusive of the constraints description and the future prospects, are summarized.

II. LITERATURE REVIEW

Neurological diseases greatly impact the population; a substantial group of the population is affected by such diseases. Various modalities are present to shed light on the insight of these disorders for diagnosing and making conclusions. Amongst these modalities, brain signals play an essential role in interpreting cognitive functionality. Supporting the researchers in discovering many defective conditions relating to specific disorders. These findings assist doctors and researchers in making proper decisions and diagnosing the disorders. Many researchers have applied computational techniques for classifying and analyzing neurological disorders.

In one of the recent efforts Ozlem et. al. proposed a methodology for detecting seizures by employing

Synchrosqueezing Transform (SST) for high-resolution time-frequency presentation on two datasets. Namely, the IKCU dataset and CHB-MIT dataset from which they have formed features like grey-level co-occurrence matrix (GLCM) and higher-order joint TF (HOJ-TF) moments. The calculated features are then passed to various machine learning classifiers achieving 95.1% accuracy, 95.54% recall and 96.87% precision for the IKCU and 95.13% accuracy, 90.30% recall and 93.37% precision [19]. Alternatively, Aayesha et al. design a methodology for classifying ictal and interictal seizure from the unknown. Taking into account two benchmark datasets Bonn and Children's Hospital of Boston-Massachusetts Institute of Technology (CHB-MIT) considers both traditional machine learning as well as Fuzzy-based approach. Inferring that both K-Nearest Neighbour (KNN) and Fuzzy rough Nearest Neighbour (FRNN) confer the best accuracy [20]. Additionally, Yong-Gi Hong et al., proposed a technique to detect the oxygen intake by the body while respiration from the nose and mouth. They assessed the EEG dataset and implemented Linear Discriminant Analysis (LDA) for feature extraction in parallel with the machine learning classifier Random Forest for classification. Discovering that nose breathing guarantees normal oxygen supply whereas breathing from the mouth disrupts the supply [21]. Utilizing the UCI-based EEG dataset Md Khurram et. al. devised a mechanism for classifying epilepsy. They applied principal component analysis (PCA) for feature extraction along with genetic algorithm (GA) for classification; they also reduced the feature set for accomplishing the task [22].

On the other hand, significant research has been published for classifying and diagnosing PD by utilizing various machine learning algorithms. Inclusive of this, Betrouni et al. in [23] proposed a methodology for classifying 100 PD patients using dense EEGs. They utilized FFT for frequency transformation and SVM and KNN as the machine learning classifier. Achieving an accuracy of 87 ± 3.5 with SVM and an accuracy of 88 ± 2.8 with KNN. Another approach has been proposed engaging QEEG that distinguishes between PD and healthy controls. The dataset consists of 79 distinct dimensions from 10 brain regions transformed to the frequency domain. The classifiers used were SVM, LR with LASSO, RF, DT, and LR, from which RF outperforms other models with an accuracy of 78% [24]. EEG has been wielded by Emad et al. for classifying PD from normal controls. They converted the resting EEGs into the deep latent space. Utilizing the orthogonalized directed coherence (gOPDC) enumerating the directional connectivity (DC) between the pairwise EEG channels among the quartet frequency (θ,α,β, and γ) and visualizing the calculated DCs into 2D plots as the input to the selected network of VGG-16 (as the pre-trained model). It achieved the accuracy of 99.62% with precision 100% and recall of 99.17% along with 0.9958 of F1- score, and the AUC averaged over 10 repetitive instances of training proffers 0.9958 [25].

A bit earlier, Shu et. al. proposed a deep learning model comprised of 13-layers that takes the pre-processed EEGs as input and gives the accuracy of 88.25% with 84.71% and 91.77% sensitivity and specificity respectively [26]. Murugappan et al. studied the emotion recognition of PD patients and normal controls (NC). Collecting the EEGs in different emotions, particularly anger, fear, sadness, disgust, and happiness. Tuneable Q wavelet transform (TQWT)is used for decomposing the signals. Eleven statistical features were extracted, and six were selected and fed to six different machine learning classifiers. From the six classifiers best performance was delivered by the probabilistic neural network with maximum mean accuracy, sensitivity, and specificity of 96.16%, 97.59%, and 88.51% respectively for normal controls along with 93.88%, 96.33% and 81.67% for PD patients [27]. Marcin et. al. proposed an approach composed of a convolutional neural network (CNN) to classify the EEGs of the patient among the three classes of Parkinsonism. From these classes, namely, mild cognitive impairment (PD-MCI), no symptoms (PD-N), and Parkinson's Disease Dementia (PD-PDD), two types of classification were proposed. One, the classification into two-tier categorization (PD-N vs MCI vs PDD) and the second one three-tier categorization (PD-N vs PD-MCI, PD-N vs PD-PDD, and PD-MCI vs PD-PDD). Two inputs were fed to CNN in the form of raw EEGs and calculated PSDs. The accuracy rises over 50% for three-class categorization and remains between 60 —70% for two-class categorization [28].

Besides, a prolific exploration was done by Sharon Hassin-Baer et al. to indicate the biomarkers related to the early stages of Parkinson's disease. Collecting the EEGs of patients as well as normal controls and calculating their event-related potentials (ERP) along with brain network analytics (BNA). They applied the machine learning mechanism to obtain a 15 BNA features-based neuro marker to differentiate the PD patient from the healthy controls [29]. Additionally, Ruilin Zhang et al. proposed a methodology to differentiate between REM sleep disorder, PD, and PD with REM disorder from the normal control for the dataset collected from Shaanxi Provincial Peoples Hospital. Two types of classification were intended by employing tuneable Q-factor wavelet transform together with deep residual shrinkage network (TQWT-DRSN) in the first experiment. Whereas wavelet packet transforms along with deep residual shrinkage network (WPT-DRSN) for the second one. Achieving 99.92% accuracy in two-class classification for PD prognosis, among the 3-class and 4-class classification tasks WPT -DRSN outperforms TQWT -DRSN with the accuracy of 97.81% and 92.59% in comparison with 95.20% and 90.46% [30].

Alongside, Sugden et al. proposed a methodology for classifying PD from the normal controls containing two datasets, specifically, the University of Mexico and the University of Iowa datasets. They applied channel-wise 1D-CNN that is trained with subject-wise cross-validation.

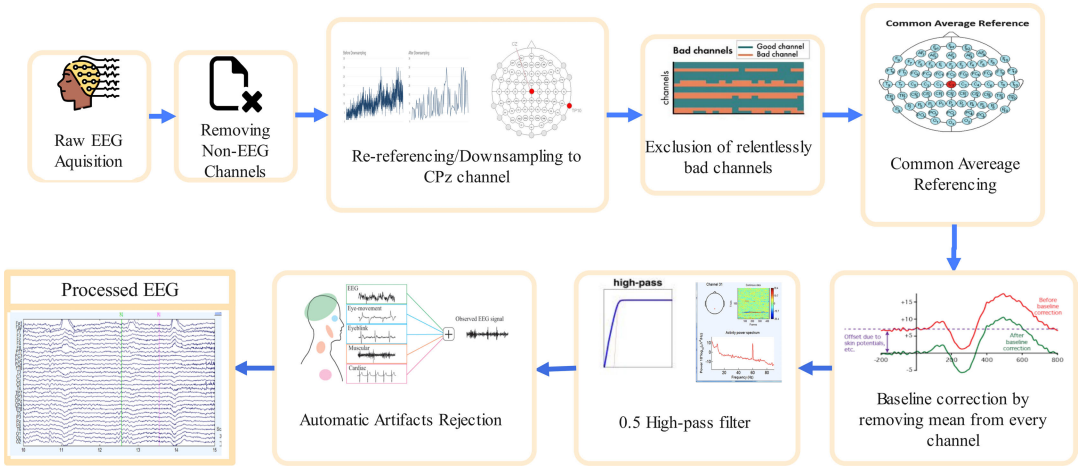


FIGURE 1. EEG preprocessing pipeline.

Their model accomplished 69.2%, 66.5%, and 72.2% of accuracy, sensitivity, and specificity, respectively, epoch-wise performance. Whereas 77.2%, 83.5%, and 71.0% of accuracy, sensitivity, and specificity subject-wise performance [31]. In another attempt, Shaban et al. proposed a 20 layered CNN architecture for classifying the PD-on medication from PD-Off medication and the healthy controls. Continuous wavelet transform (CWT) was used for frequency transformation, whereas gradient-weighted class activation mapping (Grad-Map) for visualizing the features. Their approach achieved the maximum values for evaluation metrics up to 99.9% [32]. Furthermore, Chu et al. prompted an approach for early detection of PD using EEG micro-states and deep learning. They utilized gradient-weighted class activation to depict the brain's activated regions corresponding to every micro-state group. The proposed approach produced the identification rate from 90%–99.0% [33].

With bounteous research on board, researchers had carved the way for clinicians to diagnose PD on multiple angles specifically related to EEG signals. Much research has been published considering time-frequency representation rather than on raw signals, even raw signals with a bit of pre-processing that dwell the path for real-time diagnosis. We have tried to classify the PD from normal controls without any frequency transformation. With the aim of diagnosing PD with raw EEG signals. The proposed and compared models are listed in Tables 3 and 4.

III. METHODOLOGY

A. DATA ACQUISITION

We have collected two datasets, namely, Predict Resting State Parkinson disease EEG dataset collected from an open-source depository named ''Predict'' coded with d002 taken from the paper [34]. The dataset comprises 28 Parkinson's patients with the same matched controls. The data set is recorded with the patients visiting the lab twice. Once with on-medication recording means with the prescribed medicine taken by the patient, and once without taking medicine with the time difference of 15 hours night long in between the two recordings. Also, the recordings are performed in the state of eyes open, and eyes closed with 64 EEG electrodes. Along with one VEOG channel known as 65 and the rest of three were XYZ form 66 − 68 on hand accelerometer having 500 Hz as a sampling rate, the details description is given in [34].

The second dataset is acquired from OpenNeuro, the UC San Diego dataset, which is the resting state dataset of PD patients. The dataset consists of 15 PD patients and 16 healthy controls. The recordings were taken in two stages, once on-medication and off-medication scenarios. Inbrain'sf-medication case, the patients were advised to pause their medications for 12 hours before taking the recordings. These recordings were captured using the BioSemi ActiveTwo technique with 32 channels at the sampling rate of 512 Hz, along with eight EXG channels. This dataset is abbreviated as the UCSD dataset, fully demonstrated in [35].

B. SIGNAL PROCESSING

After the EEG has been recorded, it should undergo baseline correction. It is essential due to the deviation produced by static charges, hydrated skin, and skin potentials in the EEG signals. These factors can produce an offset in EEG that gradually moves upward and downward, deviating from the original signal. Baseline corrections are used to overcome these drifts or deviations in the EEG [36], [37].

Apart from baseline correction, another deviation is produced by the noise present in the EEG signal. These noises or artifacts are the external elements arising due to various reasons explicitly, having environmental, experimental, ''nd phys''ological effects [38]. These artifacts can be categorized into extrinsic aParkinson'sc artifacts. Deliberately, extrinsic artifacts are the noise appended due to external elements like environmental errors or experimental errors. The intrinsic artifacts are associated with physiological activity (like eye

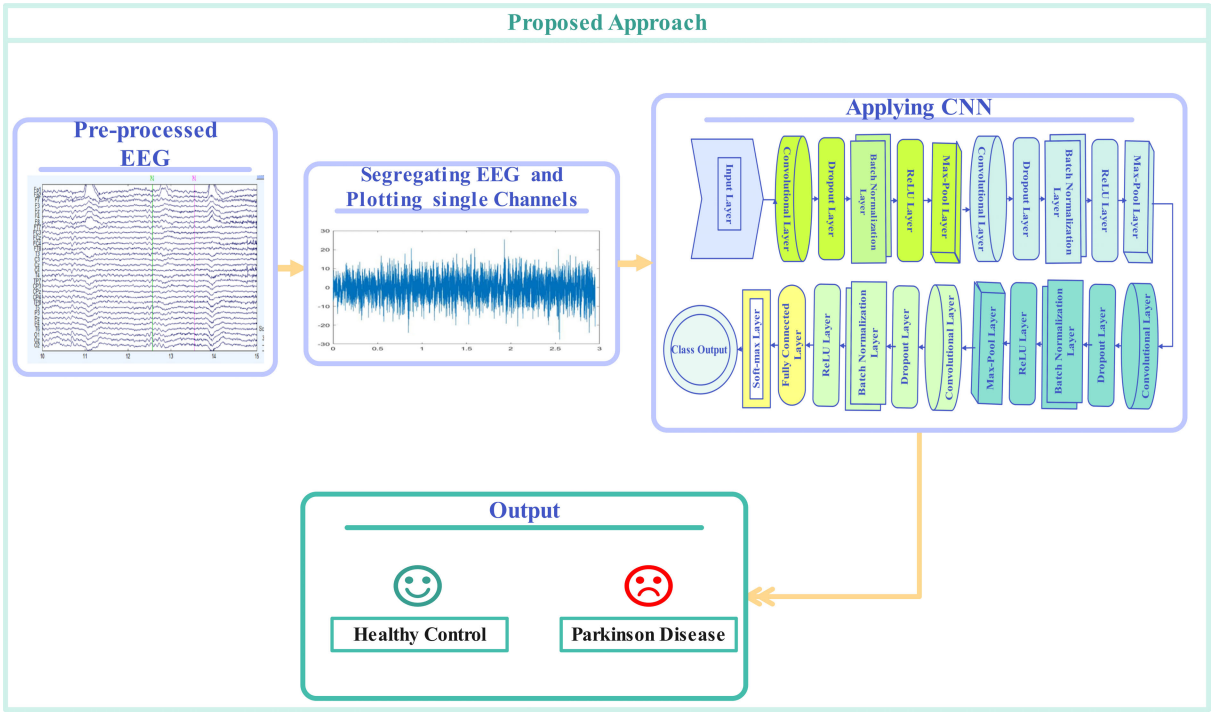


FIGURE 2. Proposed U^{EN}-PDNet.

movement, muscular activity, cardiac vibrations) [38], [39]. To deal with these noises, different pre-processing pipelines are present. The pre-processing pipelines are used to reject these artifacts and approximate EEG, making it clearer and readable.

We have followed the pre-processing steps from [34] and [35]. The steps involved in basic pre-processing are done in the same fashion. For both datasets, specific steps are the same, but there are some differences as the latter dataset contains only 32 channels, and the former contains 64. The primary EEG pipeline is described in Figure 1. General steps are described here as in

- Eradicating the VEOG channels that are non-EEG channels from the predicted dataset along with the XYZ accelerometer. In the case of the UCSD dataset EXG, eight channels should be removed.
- Re-referencing and downsampling to the central electrode known as CPz in the case of the UNM dataset and Cz in the case of the UCSD dataset.
- Eliminating the persistently lousy channels
- Excluding the contaminated channels re-referencing the channels to average.
- Removing the mean from every channel.
- For eliminating the low-frequency range, a high-pass filter is applied of 0.5 Hz (using a two-way FIR filter).
- Using EEGLAB eliminating artifacts.

For assessment, we have divided the recordings into the number of channels. Hence in the case of the UNM dataset, as each recording for the subject contains almost 60 channels, not the same for every recording, we have divided each channel into a single data point. And in the case of the UCSD dataset, there are 32 channels from which approximately 30 channels were functional due to the rejection of wrong channels. Several channels arise due to the rejection of wrong channels not persistent for every subject.

C. PROPOSED APPROACH

Subsequently, we have tried to produce significant results from the raw data. We have directly plotted the segregated EEG channels produced from the primary pre-processing pipeline using MATLAB. These plots are then fed to the 2D-CNN to make a classification. The basic outline of the proposed methodology is shown in Figure 2.

1) CONVOLUTIONAL NEURAL NETWORK (CNN)

A convolutional Neural Network (CNN) is a neural network that tries to catch the trends present in the data. Hence it is more specifically useful for pictorial or graphical representations. As these graphical representations are stored in the system digitally in the forms of grids, these networks read the inputs in the forms of grids inspired by the animal visual cortex [40]. The neurons attached to every layer function in a way that human brains perceive. The neurons in the brain react to receptive fields so that the associated neurons convey their outputs to the next layer to completely visualize the patterns. Under the human brain, every neuron responds to stimuli particular to the confined range of the visual realm, known as a receptive field in the human

TABLE 1. CNN layers.

Layer	Activation	Learnable Parameters	Number of Learnable Parameters
Input Image	500(S) × 500(S) × 1(C) × 1(B)	-	0
Convolution_2D_1	500(S) × 500(S) ×32(C) ×1(B)	Weights 4 × 4 × 1 × 32 Bias 1 × 1 × 32	544
Dropout_1	500(S) ×500(S) ×32(C) ×1(B)	-	0
Batch Normalization_1	500(S) ×500(S) ×32(C) ×1(B)	Offset 1 × 1 × 32 Scale 1 × 1 × 32	64
ReLU_1	500(S) ×500(S) ×32(C) ×1(B)	-	0
Max Pooling_1	250(S) ×250(S) ×32(C) ×1(B)	-	0
Convolution_2D_2	250(S) ×250(S) ×64(C) ×1(B)	Weights 4 × 4 × 3264 Bias 1 × 1 × 64	32832
Dropout_2	250(S) ×250(S) ×64(C) ×1(B)	-	0
Batch Normalization_2	250(S) ×250(S) ×64(C) ×1(B)	OffSet 1 × 1 × 64 Scale 1 × 1 × 64	128
ReLU_2	250(S) ×250(S) ×64(C) ×1(B)	-	0
Max Pooling_2	125(S) ×125(S) ×64(C) ×1(B)	-	0
Convolution_2D_3	125(S) ×125(S) ×128(C) ×1(B)	Weights 4 × 4 × 64 × 128 Bias 1 × 1 × 128	131200
Dropout_3	125(S) ×125(S) ×128(C) ×1(B)	-	0
Batch Normalization_3	125(S) ×125(S) ×128(C) ×1(B)	OffSet 1 × 1 × 128 Scale 1 × 1 × 128	256
ReLU_3	125(S) ×125(S) ×128(C) ×1(B)	-	0
Max Pooling_3	62(S) ×62(S) ×128(C) ×1(B)	-	0
Convolution_2D_4	62(S) ×62(S) ×256(C) ×1(B)	Weights 4 × 4 × 128 × 256 Bias 1 × 1 × 256	524544
Dropout_4	62(S) ×62(S) ×256(C) ×1(B)	-	0
Batch Normalization_4	62(S) ×62(S) ×256(C) ×1(B)	Offset 1 × 1 × 256 Scale 1 × 1 × 256	512
ReLU_4	62(S) ×62(S) ×256(C) ×1(B)	-	0
Fully Connected	1(S) ×1(S) ×2(C) ×1(B)	Weights 2 × 984064 Bias 2 × 1	1968130
Softmax	1(S) ×1(S) ×2(C) ×1(B)	-	0
Classification	1(S) ×1(S) ×2(C) ×1(B)	-	0

visionary system. Every neuron in CNN reacts similarly and processes the input in its receptive field only [41]. CNN consists of several layers, namely convolutional, pooling, and fully connected layers, with some additional layers included later on, specifically the Batch normalization layer, Dropout Layers. The Convolutional layer is the primary layer of CNN which primarily extracts the features as activations. The basic functionality of the layers is enlightened concisely in Equation (1).

$$A * B = \int C(a, b)D(x + a, y + b). \tag{1}$$

We have performed parameters tuning extensively discussed in section III-F. The initialization of the hyperparameters is done in initializing the model. The basic layer of CNN is the convolutional layer that produces the feature map by convoluting the filter size by the given image size given by the Equation (1). We have used 4 convolutional layers, and in each layer, we have doubled the number of filters starting from 32—256 inherited from the encoder part of the U-Net architecture. The dropout layer follows the convolutional layer, and the dropout layer is used to overcome the problem of overfitting. When the model is trained well but under-performs on testing or new data. We have used 0.21 as the dropout rate in the proposed model.

The dropout layer prevents overfitting, but for the non-normalized data, there is a possibility of biasedness towards higher values. To overcome this biasedness, we have used a Batch normalization layer that maintains all the outcomes in a standard scale without compromising the quality, depicted by the Equations (2), (3), (4) and (5). We have used the Batch Normalization layer with every convolutional layer to train the model correctly.

$$\bar{y}_i = \frac{y_i - \mu_a}{\sqrt{\eta_a^2 + \epsilon}} \tag{2}$$

where,

$$\mu_a = \frac{1}{n}\sum_{j=1}^{n} y_i, \tag{3}$$

is the batch mean, and

$$\eta_a^2 = \frac{1}{n}\sum_{j=1}^{n}(y_i - \mu_a)^2 \tag{4}$$

is the batch variance, hence the scaled and shifted activations are given by the equation

$$z_i = \alpha\bar{y}_i + \theta \tag{5}$$

where α and θ are the learnable parameters.

In addition, the ReLU activation function is used, followed by the Max-Pooling operation. The ReLU function increases the non-linearity for proper feature extraction expressed by

the Equation (6). Finally, the dense layer and softmax layers are used, followed by the classification layer

$$ReLU(x) = \max(0, x) \tag{6}$$

2) PERFORMANCE EVALUATION METRICES

We have used four evaluation matrices, specifically Accuracy, Precision, Recall, and F1-scores given by the Equations (7), (8), (9) and (10) respectively. Explicitly, accuracy informs us about the closeness of the predicted value to the actual value. Precision tells us how much the model is correct in optimistic predictions. The Recall, the true positive rate (TPR) or true positives, calculates the correct prediction rate. F1-score is the harmonic mean of precision and Recall which tells us about the model's performance. A more explicit stance of the terms True Positive, True Negative, False Positive, and False Negative are needed for more apparent concepts on the said matrices.

- True Positive (tp): The right positive predictions by the model.
- True Negative (tn): The right negative prediction by the model.
- False Positive (fp): The wrong positive prediction by the model.
- False Negative (fn): The wrong negative prediction by the model.

Now, the evaluation of the matrices is calculated by the formulas

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{7}$$

$$Precision = \frac{tp}{tp + fp} \tag{8}$$

$$Recall = \frac{tp}{tp + fn} \tag{9}$$

$$F1 - Score = 2 \times \left(\frac{Precision * Recall}{Precision + Recall}\right) \tag{10}$$

D. EXPERIMENTAL SETUP AND EXPERIMENTATION

We have already pre-processed the EEGs in section III-B and separated the channels into a singular window. All these signals are then plotted using MATLAB. We have also utilized the Convolutional Neural Network (CNN) for classification comprised of 23 layers inherited from the encoder part of U-Net architecture. The proposed model comprises of Convolutional layer followed by the dropout layer, batch-normalization layer, and ReLU activation along with the max-pooling layer shown in Fig 3 and Table 1. Four Convolutional, Dropout, Batch-Normalization, and ReLU Layers are used in addition to three Max-Pool layers followed by a fully-connected or dense layer, softmax, and classification layer comprised of the whole CNN model.

Inclusive of all in our experimentation, there are 6 cases for both Datasets since we have six patient cases compared to the two control cases, which can be observed from Table 2.

For every case shown in the Table, 2 number of plots differs for both the classes On-medication and Off-medication.

TABLE 2. Individual cases in both the dataset.

Patients and Controls Considered Conditions		
	Patients	**Healthy Controls**
UNM Dataset		
1)	On-Med Eyes Closed	Eyes Closed
2)	On-Med Eyes Open	
3)	Off-Med Eyes Closed	Eyes Open
4)	Off-Med Eyes Open	
UCSD Dataset		
5)	On-Med	Healthy Controls
6)	Off-Med	

In the case of On-med eyes-closed vs. Eyes Closed Controls, the image set of patients contains 1617 plots, whereas the controls' image set consists of 1398 plots. Similarly, in all the cases, the number of plots is different due to the rejection of channels and removal of non-EEG channels. The original image has the size of 875 × 656 pixels, which we have trimmed down to 500 × 500 pixels. We have divided the 70% of the Dataset for training, and from the remaining 30%, one-half is for validation, and the other half is left for testing.

E. HYPER-PARAMETERS OPTIMIZATION

Hyper-parameter optimization is an essential step in training CNN. Numerous methods are present for optimizing the hyper-parameters like Grid Search, Random Search, and Bayesian Optimization. We have used Bayesian Optimization and the hyperparameters function in MATLAB. We have already explained hyperparameters used in detail in Section III-F. We have tried both manual Optimization and Bayesian Optimization too.

In manual hyperparameters optimization, we have trained the model with different hyperparameters combinations finding the optimal model. The same is done with Bayesian Optimization with the same hyperparameter combination. The hyperparameters we have selected are the Initial Learning rate, Squared Gradient Decay Factor, Minimum batch size, and Maximum Epochs from the training options. Whereas the hyperparameters used in the layers option are Filter size for the convolutional layer, Dropout factor in the Dropout Layer, and Pooling Layer, the Pool size and Stride as the hyperparameters. In some cases, manual Optimization produced the best result, and in most cases, Bayesian Optimization produced optimal results.

However, Bayesian Optimization has converted the lengthy task into a simple one. We have tried all the possible combinations of variables, such as in the case of the Initial Leaning Rate, the tuning is done from 0.01—0.0001. The Squared Gradient Decay Factor is tested from 0.75 —0.99 for the ADAM optimizer. The Learn Rate Drop Factor is variated from 0.01—0.1, and the Learn Rate Drop Period is tested in 5 —20. Mini Batch Size is tested between 1—8, but

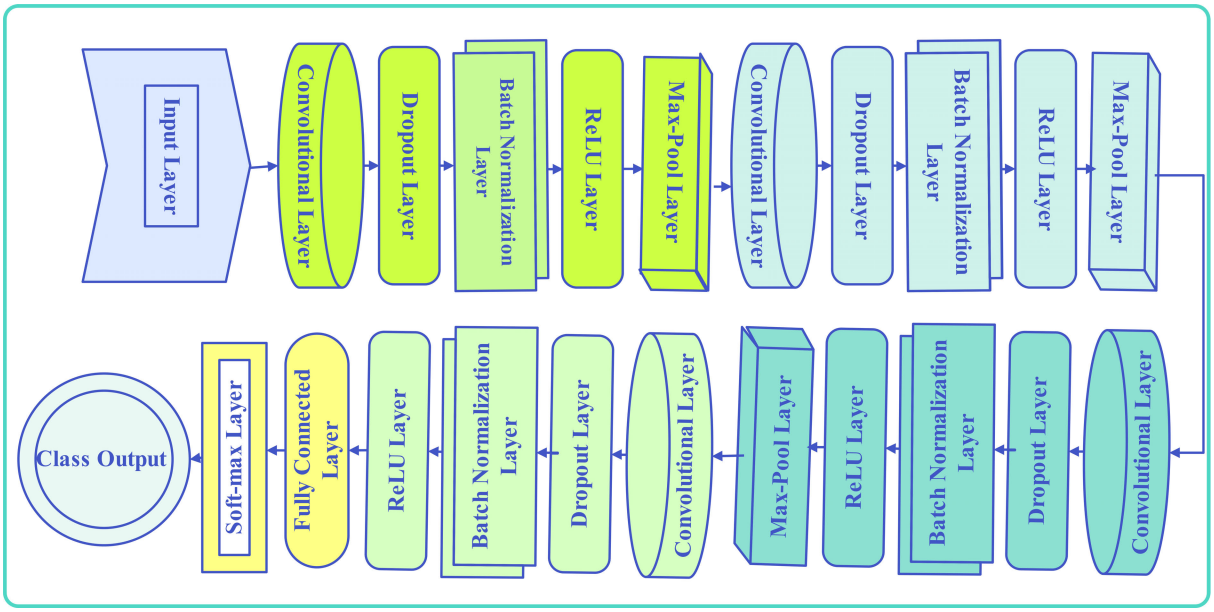


FIGURE 3. CNN layered architecture.

due to lack of resources maximum length of batch size is stuck to 8; on increment from 8, the GPU was going out of resources. Max Epochs has played an essential role in training the model; at the beginning, when we were doing manual hyperparameter tuning, we took Max Epochs as 30, but the model was not performing well both on validation as well as a testing dataset, with Bayesian Optimization the Max Epochs reached to 60 prospering best results in the case of On Med Eyes Closed of UNM dataset.

In the case of Layer hyperparameters, we have performed tuning for Convolutional Layer parameters where the filter size varied from 1—4 and the number of filters remained fixed to 32 for the first convolutional layer. For the second convolutional layer, the filter size was the same, but the number of filters was fixed to 64; similarly, for the third and fourth convolutional layers, the filter size was varied in the same manner, and the number of filters was fixed to 128 and 264 respectively. The dropout factor in the first Dropout Layer was tried between 0.1 —0.5, which gives 0.20 as the best dropout factor; it remains the same for all four Dropout Layers. For the Max-Pooling layers, the Pool size is varied between 1 —4 with Stride from 1 —3 giving the best pool size of 2 along with Stride 2.

This describes a case of On Med Eyes Closed off the UNM dataset where the Bayesian Optimization has delivered the optimized results. But in some cases, Bayesian Optimization has failed, and manual Optimization has provided the best results. Although we have performed both optimizations in every case, in ablation studies III-F, we have discussed only manual Optimization because the discussion is too long, and we cannot discuss every particular aspect in the paper.

F. ABLATION STUDIES

In the beginning, we were trying to make a classification by plotting all the channels on one graph as a regular EEG. But due to the shorter dataset, the problem of under-training prevails. Which produces the problem of underfitting, and the model performed abominably. To overcome the anterior issue, we have plotted all the channels individually obtained from the pre-processing pipeline for both UNM and UCSD datasets.

To obtain the best results, we have experimented with two optimizing techniques: adaptive Moment Estimation (ADAM) and stochastic gradient descent with momentum (SGDM). In some trials, SGDM outperformed ADAM, but mostly in all the cases, ADAM yielded the best results. In the neural network, we have trained and tested iteratively for every dataset containing all On-med and Off-med classes. The hyperparameters tuning one by one is done for every class considering one optimizer at a time to improve the classification accuracy. In the first trial, we used only the Convolutional, ReLU, and Max-pooling layers with three blocks connecting with the softmax and classification layers. The classification validation and testing accuracy was stuck at about 87% and 83%, respectively, for the UNM dataset, which we have considered first.

1) STOCHASTIC GRADIENT DESCENT WITH MOMENTUM (SGDM)

The gradient descent looks for the minima in the search space for every specific objective function, calculating the target function's derivative descending from the derivative's direction. Since the gradient or the derivative can move in the absurd direction, a hyper-parameter known as momentum is added in the gradient descent that tells about movements made in the past. The momentum value ranges from 0—1; normally assigned values are 0.8. 0.9 or 0.99. For the SGDM, we have started the momentum factor with 0.75 along other hyperparameters within the convolutional neural networks

TABLE 3. Comparison of performance of proposed approach on UCSD dataset with the existing researches.

Authors	Feature Analysis	Classification Technique	Classification Between	Accuracy	Precision	Recall	F1-Score
Khare et.al.2021a	Tunable Q wavelet Transform	Least Square SVM	PD_OFF vs HC	96.13%			
			PD_ON vs HC	97.65%			
SA Pozo Ruiz 2021	Waveform Shape Features, Spectral Features, Statistical Features	Gaussian Process Classifier, SVM, KNN, RF, CNN	PD_OFF vs HC	85.10%	86.60%	84.30%	0.854
			PD_ON vs HC	83.90%	84.00%	83.70%	0.838
Qiu et. al. 2022	PSD and Phase Locking value	LeNet-5	PD_OFF vs HC	89.16%			
			PD_ON vs HC	92.19%			
Proposed Approach	No	CNN	PD_OFF vs HC	97.90%	98.00%	97.87%	0.9794
			PD_ON vs HC	96.80%	96.85%	96.78%	0.9682

PD_OFF: Parkinson Patients when medication is off, PD_ON: Parkinson Patients when medication is on, HC: Healthy Controls.

(CNN). In the first iteration, we used the filter size of 4 with a 32 number of filters in the first convolutional layer. Whereas, 50 is in the second convolutional layer and 100 in the third convolutional layer. In addition, we have used the ReLU layer and max-pooling layer having filter size 2×2 with a stride of 2 too. The first iteration took 10 hours, producing an accuracy of 83% with an image input size of 500×500.

In the next recurrence, add the dropout layer with dropout factor 0.5 in addition to filter size 4 and several filters 32 for the first convolutional layer; for the second convolutional layer only the number of filters is increased to 64 with all the rest remaining the same. Similarly, in the third convolutional layer, the number of filters increased to 128 the filter size remained the same. From the first trial, after adding other layers, the testing accuracy achieved was 82% with a validation accuracy of 89%; the problem of overtraining persists, giving the network a tough time for the testing dataset. Hence, with every convolutional layer, one dropout layer, one Batch Normalization layer, one ReLU layer, and one Max-pooling layer are appended. Except with the last convolutional layer, instead of the Max-pooling layer fully-connected layer, the softmax layer and the classification layer are added, producing approximately the same results as in the previous.

It was observed that not any beneficial results were obtained from the previous trials. Henceforth, we increased the count of every Convolutional, Dropout, Batch normalization, ReLU, and Maxpooling layer by one. With momentum set to 0.75, the filter size of the convolutional layer is set to 4 for all four layers, along with the number of filters set to 32, 64, 128, and 250 for the first, second, third and fourth convolutional layers respectively. The dropout factor of 0.25 remains the same for the first dropout layer, whereas it was set to 0.2 for the following three layers. In addition, the max-pooling filter size was set to 2 along with the stride of 2×2 for all four max-pooling layers. The number of epochs was increased to 60 from the 30 in the previous trial with the initial learning rate of 0.0001. It produced the validation accuracy of 91.10% and the testing accuracy of 90.80% for the off-medication eyes open cases. In contrast, the validation accuracy achieved for the eyes closed cases was 91.30%, along with the testing accuracy of 91.00%. This was the best metric achieved overall for the UNM dataset amongst all the conditions.

We continued the iterations variating different hyper-parameters; likewise, we varied the momentum to 0.80, bounding other hyperparameters to the same as previous. For all the conditions starting from the off-medication eyes open, the validation accuracy obtained was 90.64% with a testing accuracy of 90.20%. In addition, for off-medication eyes closed, the validation accuracy attained was 88.40%, and the testing accuracy of 86.50%. For medication, it remains approximately the same. Whereas in the case of on-medication eyes closed, the validation accuracy reached 89.00% with testing accuracy of 85.00%. Also, we had performed intensive experimentation with umpteen iterations for varying the hyperparameters, but the best results achieved were in the case of 0.75 momentum for off-medication cases. On the other hand, in the case of on-medication, the best results were attained with 0.75 momentum and the dropout factor of 0.21 for the first layer and 0.20 for the rest of the layers.

2) ADAPTIVE MOMENT ESTIMATION OPTIMIZATION (ADAM)

Adam could be recognized as combining Root Mean Square Propagation (RMSProp) and the SGDM since it utilizes the squared gradients for scaling the learning rate imported from the RMSProp along with the momentum incorporating

TABLE 4. Comparison of performance of proposed approach on UNM dataset with the recent researchers.

Authors	Feature Analysis	Classification Technique	Classification Between	Accuracy	Precision	Recall	F1-Score
Anjum et. al. (2020)	Linear Predictive Coding	Hyper Planes	PD_EO vs HC_EO	78.70%			
			PD_EC vs HC_EC	82.20%			
SA Pozo Ruiz 2021	Waveform Shape Features, Spectral Feature, Statistical Features	Gaussian Process Classifier, SVM, KNN, RF, CNN	PD_OFF vs HC	87.70%	87.70%	87.40%	0.875
			PD_ON vs HC	86.90%	90.00%	83.00%	0.863
Lee et. al. 2022	Hjorth Parameter	Gradient Boosting Decision Tree	PD vs HC	89.30%			0.903
Shah et. al. 2022	Discrete Wavelet Transform (DWT), Differential Entropy and Hjorth Parameters	KNN	PD vs HC	88.51%	89.22%	91.58%	
Proposed Approach	No	CNN	PD_ON_EC vs HC_EC	91.00%	91.50%	90.68%	0.9109
			PD_ON_EO vs HC_EO	91.20%	91.23%	91.17%	0.912
			PD_OFF_EC vs HC_EC	91.50%	91.70%	90.31%	0.9151
			PD_OFF_EO vs HC_EO	93.10%	93.18%	93.09%	0.9313

PD_OFF: Parkinson Patients when medication is off, PD_ON: Parkinson Patients when medication is on,
HC: Healthy Controls, PD_ON_EC: Parkinson Patients with medication ON and Eyes Closed,
PD_OFF_EO: Parkinson Patients with medication OFF and Eyes Open,
HC_EC: Healthy Controls with Eyes Closed, HC_EO: Healthy Controls with Eyes Open

the moving average of the gradient instead of the gradient analogous to SGDM.

On the experimentation part, we took all the conditions one by one to improve the model's performance. Inclusive of the results obtained as in the case of SGDM optimizer, we tested all the conditions with Adam to vary the squared gradient decay factor set to 0.80 and other hyperparameters for both the datasets. Starting from the UNM dataset with Off Medication eyes closed condition, the experimental results obtained were almost similar compared to the SGDM optimizer in preliminary steps. Setting the filter size to 2 and the number of filters to 30 for the first convolutional layer with the same filter size with several filters were set to 60 and 100 in the second and third convolutional layers. The dropout factor was also set to 0.40 for the first trial with a filter size of 2 and the stride of size 2 too in all the Maxpooling layers.

From our experience with the SGDM optimizer, the dropout factor between 0.20—0.25 produced good results. Hence, in these cases, we maintained the dropout factor in this range only. In preliminary steps, we considered three convolutional layers, three Dropout layers, three Batch Normalization layers, three ReLU layers, and two Maxpooling layers connected with one Fully connected layer, one Softmax layer, and one Output layer. Image size remains 500×500 with a dropout factor of 0.25 for each layer. The filter size is set to 4 for all the convolutional layers alongside 30, 60, and 100 filters in the three layers, respectively. The maximal accuracy achieved in this case was 87.4% for validation with a testing accuracy of 83.6%.

Since the evaluation metrics did not provide improved results, in the next iteration, the dropout factor for the first layer was dropped to 0.20 with 0.25 in the next two dropout layers. The gradient decay factor was shifted to 0.90, and the filter size was the same in this trial, with the number of filters increased to 32 in the first convolutional layer along with 64 and 128 for the second and third layers. In the Maxpooling 2D layer, the filter size was also set to 2 with stride 2. The validation accuracy achieved in this case was 90.70% with a testing accuracy of 84.70%. Although the training was relatively successful in testing the dataset, the results showed degraded trends.

In the search for better results, we performed the subsequent trial by modifying the hyperparameters. The results had shown the number of filters worked well in the previous setup for the convolutional layers in the order of 32, 64, and 128. Also, we tried the dropout factor varying from 0.4—0.2 with the max-pooling filter of size 2. In this iteration, we updated the dropout factor to 0.21 for all three dropout layers with several filters set to 100 for the third convolutional layer. To get improved results, the image size was reset to 567×500 with the caution of not losing the necessary information. The filter size of the Maxpooling layer remains 2 with the stride 2. The squared gradient decay factor was set to 0.90 in the training options. The output attained was

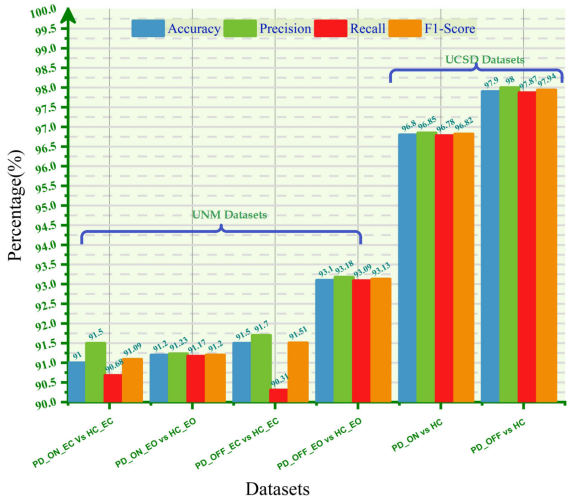


FIGURE 4. Evaluation metrics of datasets.

not commendable. Instead, the result drowned again with a validation accuracy of 87.70% and the testing accuracy of 85.20%. It is clear from this trial that increasing the image size doesn't improve the results. We had performed numerous iterations with increasing image size but showed only the best one.

In the next attempt, every Convolutional layer, Dropout Layer, Batch Normalization layer, ReLU layer, and Max-pooling layer was increased by one. The filter size of the convolutional layer remains 4 for all four layers just the number of filters was set to 32, 64, 128, and 256 for the four layers, respectively. The dropout factor was set to 0.21 for all four layers, including a filter size of 2 for the Max pooling layer with a stride of 2 for all four layers. The initial learning rate was set to 0.0001 with the squared gradient decay factor setting to 0.99. The input image size was reset to 500 × 500; for the last several experiments, it was 567 × 500. The results were quite moderate, with a validation accuracy of 91.14% and the testing accuracy of 89.30%.

The input image size was variated between 800 × 656 to 500 × 500, which increases the training time. In addition, the accuracy remains the same. Hence, the image size was reset to 500 × 500 to improve the time complexity and accuracy. With such deviation in the accuracy, we have performed many hyperparameters tuning inclusive of the number of epochs, the number of filters in the convolutional layer, the size of filters, the dropout factor, image size, and the number of layers.

Finally, we got some valuable results for the testing accuracy. The final configuration of the hyperparameters was primarily the input image size reset to 500 × 500. For the convolutional layers, the filter size was 4 for all four layers, and the number of filters was 32, 64, 128, and 264. The dropout factor finally used was 0.21 for all four layers. The filter size of 2 with the stride of 2 was set for the max pooling layer. The number of epochs was tested between the range 30—60 with best results achieved at 60 epochs. With such extensive effort, we have configured the hyperparameters taking the maximal accuracy of 93.10% for the instance of PD_OFF_EO vs HC_EO in the UNM Dataset. On the other hand, the optimal accuracy achieved of 97.90% for PD_OFF vs HC case of UCSD Dataset produces an enhanced performance with other evaluation metrics. This procedure was done for all the datasets described above, producing different results. The final results are depicted in Figures 4 and 5.

IV. RESULTS

We have proposed the CNN model for dataset sub-divisions mentioned in Table 2. Totally 6 experimentation was concluded for both datasets; in particular, as in the case of the UNM dataset, 4 significant conditions exist. Exclusively, On medication, eyes open (PD_ON_EO) vs. eyes open, healthy controls (HC_EO), On medication, eyes closed (PD_ON_EC) vs. eyes closed healthy controls (HC_EC), Off medication eyes open (PD_OFF_EO) vs. eyes open, healthy controls (HC_EO) and Off medication eyes closed (PD_OFF_EC) vs. eyes closed healthy controls (HC_EC). Whereas, for the UCSD dataset, two conditions exist specifically, On medication (PD_ON) vs. healthy controls (HC) and Off medication (PD_OFF) vs. healthy controls (HC) conditions. The proposed model is unique in that it is a derivative of U-net architecture with only encoder properties; it is also operating on the raw data. The results show the proposed methodology is an improvement in analyzing the data in raw format.

In the case of the UCSD dataset, two conditions exist, On medication and Off medication. The proposed model has illustrated that the results obtained from the discussed approaches considering the same dataset are enhanced. Inclusive of, [41] in which Khare et al. decompose the EEG into sub-bands using an automated tunable Q wavelet transform (A-TQWT) and finally using a least square support vector machine (LSSVM) for classification. A classification of PD_OFF vs. HC achieved an accuracy of 96.13% along with PD_ON vs. HC with a classification accuracy of 97.65%. In [42], Ruiz proposed a technique for classifying Parkinson patients with healthy controls by considering both UNM and UCSD datasets. They had classified with multiple classifiers amongst which SVM outperforms others with accuracy, precision, recall, and F1-score of 85.10%, 86.60%, 84.30% and 0.8540 for PD_OFF vs. HC respectively. On the other hand, for PD_ON vs HC the accuracy, precision, recall and F1-score 83.90%, 84.0%, 83.70% and 0.8380 respectively. Another methodology Qiu et al. proposed utilizes the power spectral densities (PSD) and phase-locked values (PLV) for assessing the multi-pattern analysis and considering both the group on medication and off medication by utilizing LeNet-5 as the classification algorithm achieving the classification accuracy of 89.16% for the PD_OFF vs HC and 92.10% for the PD_ON vs. HC [43].

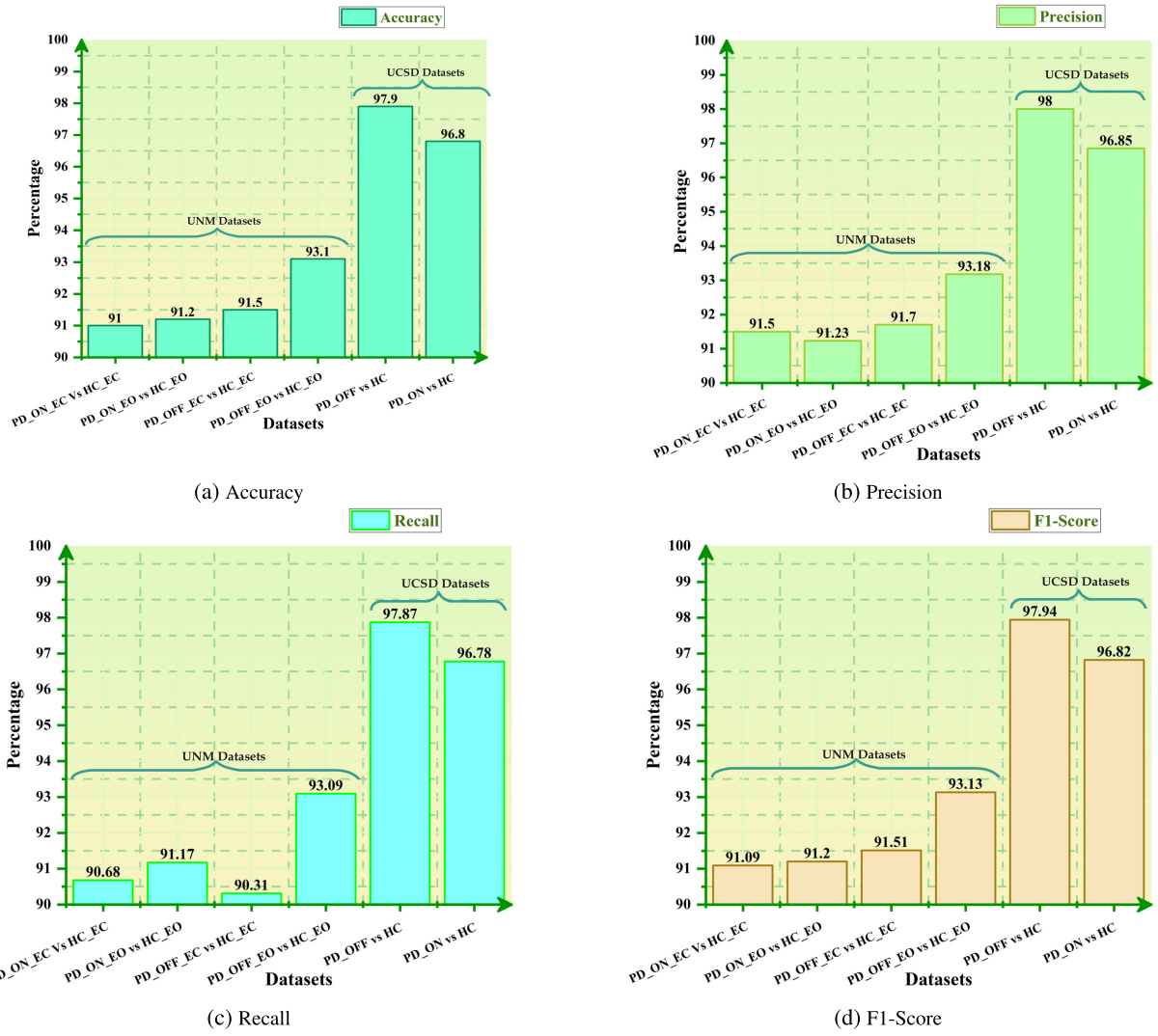
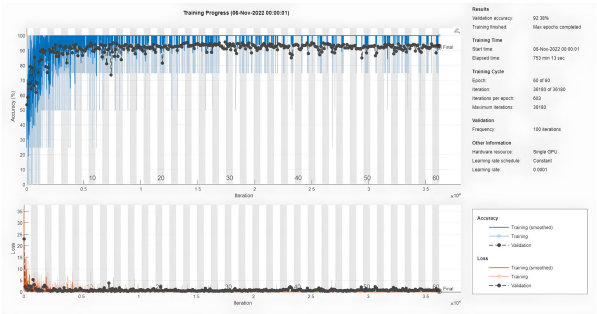


FIGURE 5. Individual evaluation metrices comparison.

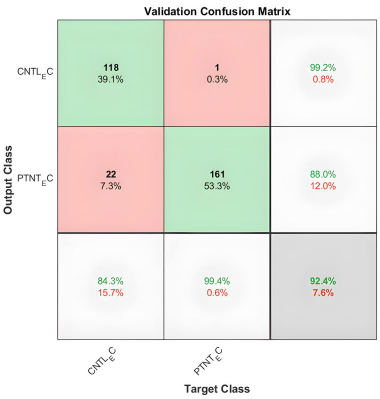
The proposed approach outperforms the discussed methods by applying the generic methodology comprising plots of EEGs fed to the 23 layered CNN. Achieving the classification testing class accuracy of 96.80% for the PD_ON vs. HC in addition to precision, recall, and F1-score of 96.85%, 96.78%, and 0.9682, respectively. They are coupled with 97.9% accuracy for PD_OFF vs. HC and precision-recall and F1-score of 98%, 97.87% and 0.9794, respectively. Table 3 depicts the details of the UCSD dataset.

For the UNM dataset, the EEG recording contains the conditions of Off medication with eyes closed and eyes open and On medication with eyes open and closed. In most of the discussed approaches, as in [44], Anjum et al. proposed a method for EEG extracting features from power spectral densities (PSD) of EEG by linear predictive coding (LPC) and incorporating hyperplanes to distinguish between the PD patients and the healthy controls accomplishing accuracy of 78.7% for PD_EO vs. HC_EO and 82.20% for PD_EC vs. HC_EC.

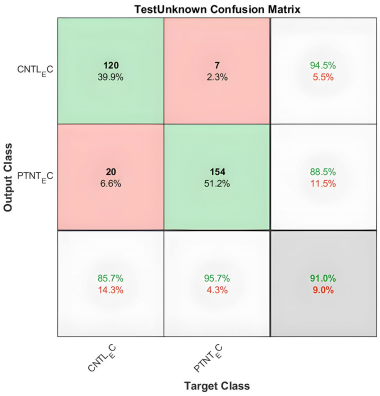
Furthermore, Ruiz in [42] implemented a rich set of features accommodating waveforms shape, spectral and statistical features. Also, for the classification of PD_OFF vs. HC as well as PD_ON vs HC, they used multiple classifiers. Amongst which SVM outperforms others, attaining the accuracy, precision, recall, and F1-score of 87.7%, 87.7%, 87.4%, and 0.875 respectively for the PD_OFF vs. HC. Together with 86.9%, 90%, 83% and 0.8630 of accuracy, precision, recall and the F1-score respectively for PD_ON vs HC. Besides, Lee et al. propounded a scheme for restricting PD patients from healthy controls. Extracting the features using Hjorth parameters and classifying using the Gradient Boosting Decision tree (GBDT) provides an accuracy of 89.30% with F1-score 0.9030 [45]. Likewise, Shah et al. formulated a scheme to differentiate PD vs Controls deducing



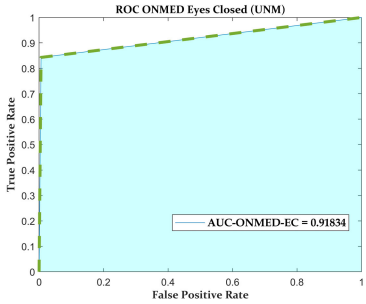
(a) Training Progress for ONMED EC Patient vs. EC Healthy Controls



(b) Validation Confusion Matrix of ONMED EC

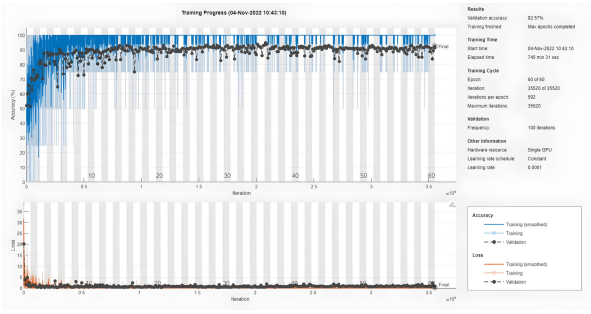


(c) Testing Confusion Matrix of ONMED EC

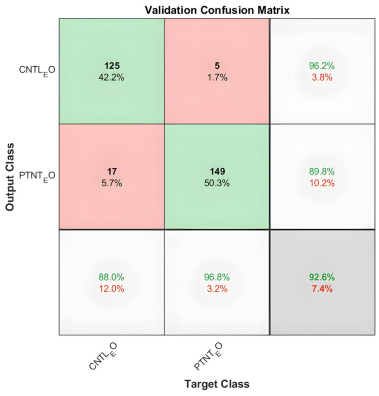


(d) ROC Curve ONMED EC

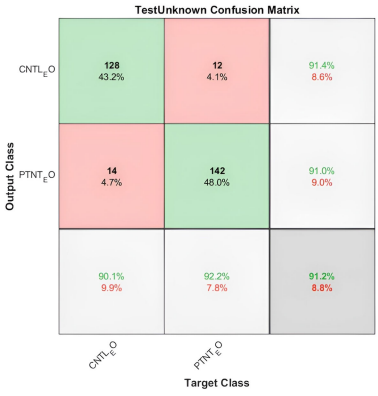
FIGURE 6. Results for ONMED EC patients vs EC healthy controls.



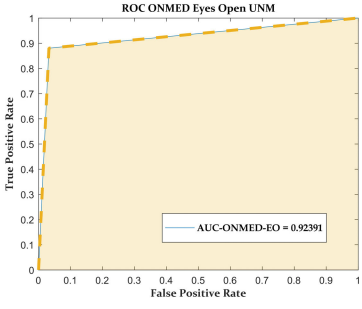
(a) Training Progress for ONMED EO Patient vs EO Healthy Controls



(b) Validation Confusion Matrix ONMED EO



(c) Testing Confusion Matrix ONMED EO



(d) ROC Curve ONMED EO

FIGURE 7. Results for ONMED EO patients vs EO healthy controls.

the features using discrete wavelength transform (DWT), Differential entropy, and Hjorth parameters. They used two classifiers CNN and KNN. The DWT-transformed images were fed to CNN, producing an accuracy of 82.5%. Whereas,

for KNN the features extracted from Differential Entropy and parameters were used, producing an accuracy, precision, and recall of 88.51%, 89.22% and 91.58% respectively [46]. On the contrary, in the proposed approach, the features are retrieved by CNN for characterizing the PD patients from the healthy controls. We have performed four different computations inclusive of on-medication and off-medication. Individually, PD_OFF_EC vs HC_EC, PD_OFF_EO vs HC_EO, PD_ON_EC vs HC_EC and PD_ON_EO vs HC_EO. Achieving the performance metrics better than the previous approaches depicted in Table 4. Besides, all the experimentation is done on Hp OMEN series laptop having the configuration, Intel(R) Core(TM) i5-7300HQ CPU 2.50 GHz processor containing one NVIDIA GeForce GTX 1050 GPU.

We have performed 6 different analyses based on the division of dataset described in Table 2. We have taken all channels with the condition's eyes closed and eyes open within the on-med and off-med conditions. The results are described here.

I.) UNM Dataset

For the UNM dataset, all four conditions displayed in Table 2 are addressed. The results achieved in all these conditions are discussed below with all the possible details.

i.) On-med Eyes Closed Patient vs Eyes Closed Control:

The results show differentiation accuracy between the patients and the control is 92.4% of validation accuracy with 91% of testing accuracy in addition to 91.5% of Precision, Recall of 90.68% and F1-score of 0.9109. Which shows clearly the emendation in the resulting metrics depicted in figures Fig 6a, Fig 6b and Fig 6c. The confusion matrix reveals the insight of the predictions made by the model; since it can be observed from the matrix, the correct predictions are shown in the diagonal that are avowed as true positives (tp) and true negatives (tn).

The matrix depicts that 120 cases are correctly predicted as Controls with eyes closed out of 301 total readings present in the testing dataset that is 39.9%, similarly 154 cases are classified as Patients with eyes closed that would be 51.2% of all the readings present. 7 patients are incorrectly predicted as Controls equate to 2.3% in consideration of all the testing sets. Comparatively, 20 of the Controls are wrongly predicted as Patients contributing 6.6% across the whole testing set.

On the whole, 91% cases are correctly predicted while 9% are wrong. The good part about the model is that it concluded with minimized type II errors. In our case, the wrong predictions made by the model as a type I error is 2.3%, which is a false positive. Hence only 2.3% of the patients are wrongly predicted as controls, whereas 6.6% among the controls are predicted as patients. Noticeably, we have described the results obtained from the testing dataset only. The validation part is shown in figure 6b.

Conclusively, it could be safer if the controls are predicted as a patient. It might not be an expensive trade-off as when the patient may be predicted as a control that might lead to disastrous outcomes. Furthermore, the evaluation metrics we have explicitly obtained, Precision of 91.5% is close to the Recall of 90.68% with the F1- score of 0.9109. Thus, results depict the efficiency of the model in the case of the On Medication Eyes Closed condition. The high and closely related values of Precision, Recall, and the F1-score demonstrate that the model performed well on the dataset apart from the Accuracy.

Also, the AUC Fig 6d depicts the correctness of the model as ROC depicts the inclination of the model towards predicting the correct instances over the incorrect instances.

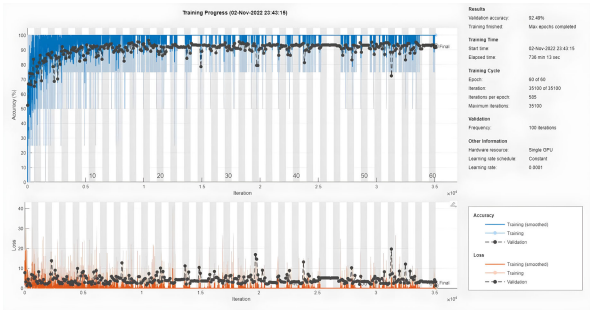
ii.) On-med Eyes Open Patient vs. Eyes Open Control:

On the Eyes Open dataset, the results show a bit different trend but approximately the same in Accuracy, which is 92.57% of validation accuracy as defined by the result shown. With the 91.2% testing accuracy in addition to 91.23% of Precision, 91.17% of recall, and the F1-score of 0.9120 as shown in figures Fig 7a, Fig 7b and Fig 7c. Also, AUC depicted in the ROC-Curve figure 7d depicts the efficacy, showing the trade-off between the actual positive rate (TPR) and the false positive rate (FPR). Similarly, the confusion matrix obtained in the On-med Eyes Open scenario reflects the results that can be observed from the diagonal of the matrix.

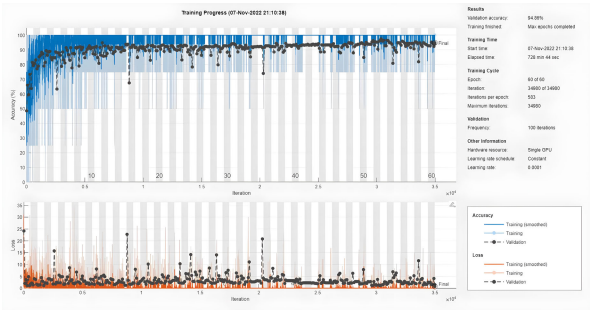
The confusion matrix shows that the true positives predicted by the model are 128 constituting 43.2% from the set of total 296 readings. Correspondingly, the true negatives predicted by the model are 142 corresponding to 48% from the set of 296 sample set. Concurrently, 12 patients are wrongly predicted as Controls compounding 4.1% whereas 14 controls are incorrectly predicted as patients contributing 4.7%.

Out of 140 Controls 128 are correctly predicted makes 91.4% whereas, 8.6% are wrongly predicted as patient. Additionally, from the set of 156, Patients' predictions 142 are precisely predicted making 91% of the total patient's predictions. On the same side, 14 are wrongly predicted as controls contributing 9% of the total patient's predictions.

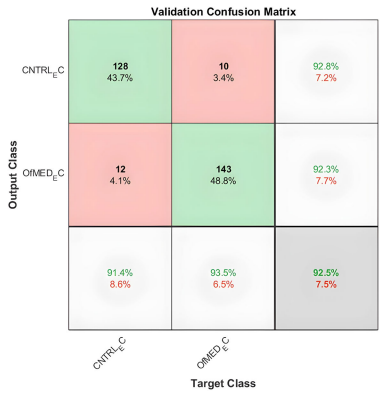
In the class of Controls eyes open 128 cases are correctly predicted that makes 90.10% of the total controls cases. While 14 cases are predicted as patients comprising 9.9%. In the case of Patients with eyes open 142 cases are correctly predicted by the model among 154 total patients that constitutes 92.2% and 12 cases are wrongly predicted as controls constituting 7.8% of the total patients with eyes open.



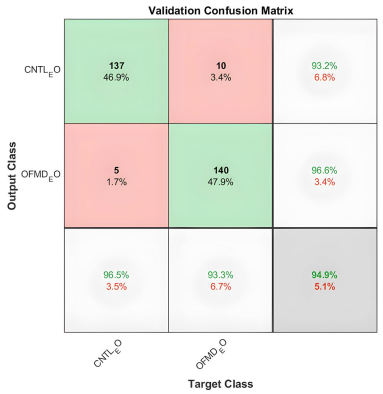
(a) Training Progress for OFFMED EC Patient vs EC Healthy Controls



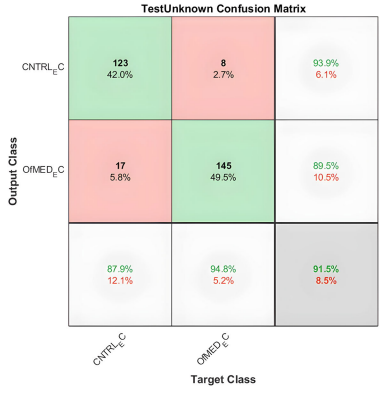
(a) Training Progress for OFFMED EO Patient vs EO Healthy Controls



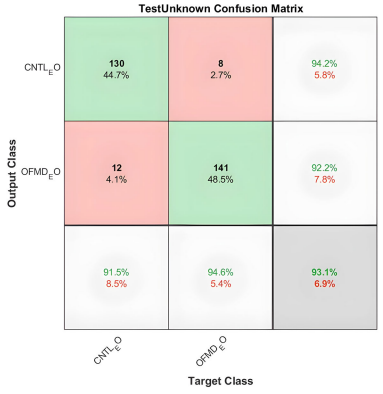
(b) Validation Confusion Matrix OFFMED EC



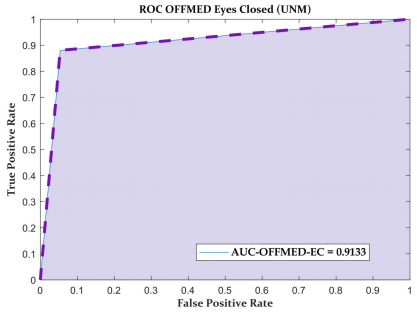
(b) Validation Confusion Matrix OFFMED EO



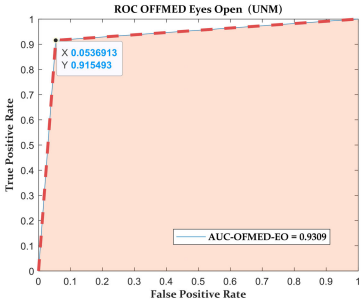
(c) Test Confusion Matrix OFFMED EC



(c) Test Confusion Matrix OFFMED EO



(d) ROC OFFMED EC



(d) ROC OFFMED EO

FIGURE 8. Results for OFFMED EC patients vs. EC healthy controls.

FIGURE 9. Results for OFFMED EO patients vs EO healthy controls.

Comprehensively, in the On med Eyes Open condition 91.2% cases are correctly predicted whereas 8.8% are wrongly predicted. Also, the type I error (False positives) in this case too is comparatively low than the

type II error (False negatives) in the testing part of the dataset.

Meanwhile, the evaluation matrices inclusive of Precision that is 91.23% is very close to Recall of 91.17%. In addition, the F1-score of 0.9120 exhibits the balance between Precision and Recall. Specifically, we have discussed the results obtained from the testing division of the dataset apart from the validation part.

iii.) Off-Med Eyes Closed Patient vs Eyes Closed Control:

In the case, off-med eyes closed, the validation accuracy results in 92.49% with the test data accuracy of 91.50% with the Precision of 91.70%, Recall of 91.31% and F1-score 0.9151. Showing an impressive improvement in comparison with the previous results can be seen in figures Fig 8a, Fig 8b, Fig 8c, and Fig 8d. The results show notable improvements also described by the confusion matrix. The confusion matrix obtained explains the model performance clearly in both the subsets of the divided dataset into validation and the testing dataset. We are discussing only the improvement shown by the confusion matrix obtained from the testing dataset. It can be observe from the confusion matrix the diagonal elements showing the correct predictions made by the model. Explicitly, 123 cases are correctly predicted as controls with eyes closed corresponds to 42% out of total 293 readings. Furthermore, 145 cases are correctly predicted as Patients off med with eyes closed that constitutes 49.5% of the total cases.

8 of the cases are incorrectly predicted as Controls with eyes closed corresponds to 2.7% among the total cases. In the same manner 17 of the cases are incorrectly predicted as Off med Patients constituting 5.7% of the total cases.

Amongst the 131 total Eyes Closed Controls predictions 123 cases are correctly predicted that is 93.9% and 6.1% of the cases are wrongly predicted. In the same manner, 89.5% of the total Eyes closed Patients predictions are correctly predicted, whereas 5.8% are wrongly predicted as the Controls.

Out of total 140 Controls readings, 123 are correctly predicted that attaining 87.9% accuracy. On the other hand, 17 cases are wrongly predicted as Eyes Closed Patients consisting of 12.1% of the total Controls cases. Among the 153 total Off med Eyes Closed Patients cases, 145 are correctly predicted as the Patients corresponds to 94.8% accuracy. On the contrary, 8 cases are incorrectly predicted as Controls making 5.2% of the total cases.

Conclusively, 91.5% are correctly predicted among the total predictions, whereas, 8.5% are incorrectly predicted. Apart from the accuracy, the type I error (False Positives) 2.7% is much lower than the type II error (False Negative) 5.8%. False positives are the cases where patients are incorrectly predicted as controls, whereas False negatives are the cases where Controls are wrongly predicted as patients. It is one of the model's positive indications, especially in our cases where if the Patient is miss-classified as the Control, it leads to a tragic outcome: False negatives 2.7% compared to 5.8% of False positives.

Along with the accuracy, Precision quantified as 91.70% in addition to the Recall calculated as 91.31% with the F-1 Score of 0.9151. The evaluation metrics show the model's performance; the F1-score is related to Precision and Recall. This closeness concludes that the model performed well in detecting the Patients in the case of Off Med Eyes Closed Patients on the testing dataset. In addition, the AUC in the ROC curve evaluates to 0.9133, confirming the model performance's efficacy.

iv.) Off-med Eyes Open Patient vs Eyes Open Control:

In the scenario of off-med Eyes open, the validation accuracy of 94.90% with testing accuracy of 93.10% in addition with Precision 93.18%, Recall of 93.09% and the F1-score having 0.9313 displayed by the figures Fig 9a, Fig 9b, Fig 9c, and Fig 9d.

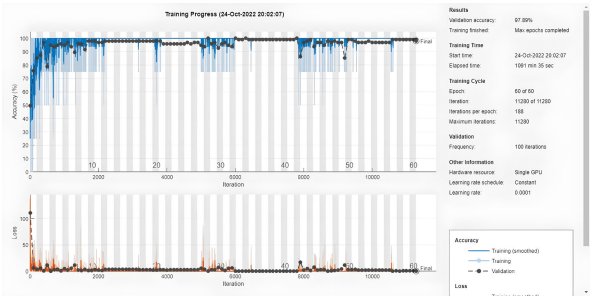
In Off medication Eyes Open instance, the results obtained are discussed earlier, specifying the values obtained for accuracy, precision, recall, and F1-score. On the other hand, we have a confusion matrix that speaks more precisely about the outcomes. The diagonal elements shown in the Fig 9c display the accurate predictions made by the model. In the current scenario, 130 cases are correctly predicted as Controls with eyes open, corresponding to 44.7% of the total 291 cases. Whereas 141 cases are accurately predicted as Patients with eyes open in the case of Off medication Eyes open, contributing 48.5% of the total cases. On the contrary, 8 cases are wrongly predicted by the model as Control constituting 2.7% of the total cases, and 12 cases are incorrectly predicted as Patients from the total cases composing 4.1% of the total cases.

Among the total 138 Control predictions, 130 cases are correctly predicted by the model, reaching the accuracy of 94.2%, whereas 8 are incorrectly predicted, that is 5.8% are wrong predictions. On the contrary, out of 153, total Patients' predictions of 141 are correctly predicted as Off-med Eyes open Patients that produce an accuracy of 92.2% and the remaining.

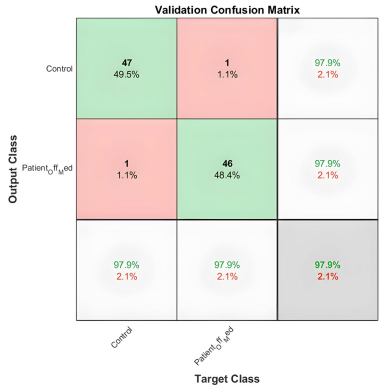
7.8% are the wrong predictions.

In the set of total 142 Controls with Eyes Open, 91.5% of the cases are correctly predicted as Controls that counts 130. Contrastingly, 12 cases are incorrectly predicted as Off-med patients imparting 4.1%. On the other hand, among the 149 Off-med Eyes closed patients, 141 correct predictions are made ensuing 94.6% of accuracy. With 8 cases are wrongly predicted as Controls that is remaining 5.4% of the total proportion.

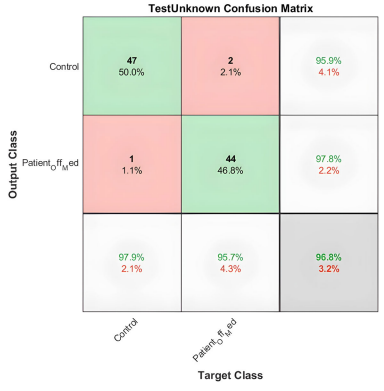
Exclusively, AUC in the ROC-curve manipulates to 0.9309 demonstrating the efficiency of the model prediction in the light of Accuracy, Precision and F1-score.



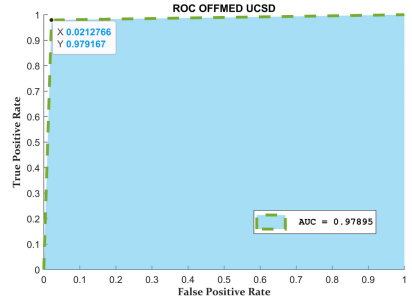
(a) Training Progress for OFFMED Patient vs Healthy Controls



(b) Validation Confusion Matrix OFFMED



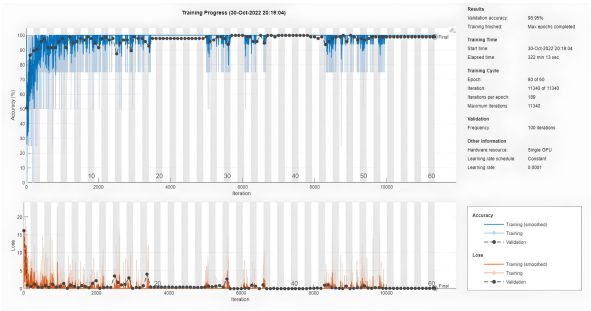
(c) Test Confusion Matrix OFFMED



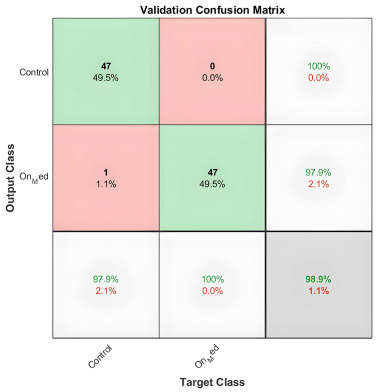
(d) ROC OFFMED

FIGURE 10. Results for OFFMED case.

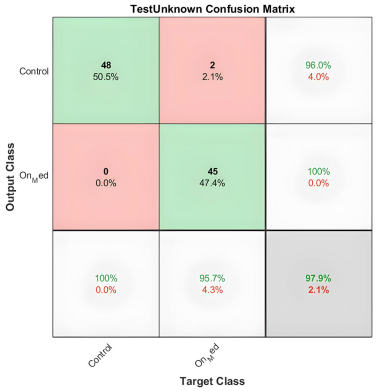
Predominantly, 93.1% of the total cases are correctly predicted containing both the Patients as well as Controls in the scenario of Off-med Eyes Closed along with 6.9% of the incorrect predictions.



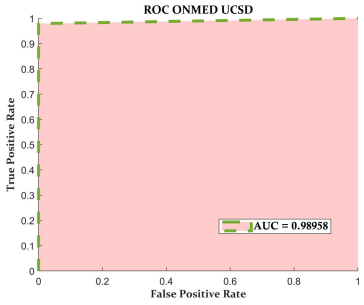
(a) Training Progress for ONMED Patient vs Healthy Controls



(b) Validation Confusion Matrix ONMED



(c) Test Confusion Matrix ONMED



(d) ROC ONMED UCSD

FIGURE 11. Results for ONMED case.

Alternatively, the evaluation matrices we have calculated depict the efficacy of the model. By producing the

Precision of 93.18%, Recall of 93.09% in addition to the F1-Score 0.9313. The yielded values demonstrate the effective performance of the model. Also, the accuracy achieved in the case of Off-med Eyes Open is the best one among the set of four instances in the UNM dataset. In addition, the Precision, Recall, and F1-Score also demonstrate the performance of the model in Off medication Eyes Open scenario.

From the results achieved in all the four considered instances of the UNM dataset are the best ones. We have tried as discussed in the III-F different model specifications to achieve the best possible evaluation matrices for all the cases independently. What we have achieved is detailed in the above sections. Amongst all the cases the best performance evaluation matrices are achieved in Off medication Eyes Open scenario.

II.) UCSD Dataset

In the UCSD dataset, there exist two situations one On medication and the second one is Off medication. The name explains the said conditions, on medication is the case when the subject is having EEG recording when the subject have taken the prescribed medicines. On the other hand, Off medication is the case when the subject didn't have taken the medicine within 12 hours of the EEG recording.

We have applied the designed model with both the variations first with the SGDM classifier along with Adam classifier alternatively. The results obtained are better than as achieved in the case of UNM dataset. The detailed results are discussed below with evaluation metrics.

v.) Off-Med vs Healthy Control:

The first condition in the UCSD dataset starts with Off-medication, the validation accuracy achieved for the Off-medication case is 97.90% along with the accuracy from the test set is 96.80%. In addition, with the Precision 96.85%, Recall of 96.78% and the F1-score 0.9682 can be observed from the figures Fig 10a, Fig 10b, Fig 10c, and 10d.

The confusion matrix is a valuable metric for displaying the efficiency of the model which cannot be achieved from other metrics. The diagonal element of the Confusion matrix demonstrates the correct predictions made by the model which is 47 cases are correctly predicted as Controls that represent true positives with 44 cases correctly predicted as Patients that is known as true negatives.

Among the total 49 Control predictions 47 are correctly predicted as controls achieving 95.9% accuracy. With 2 incorrect predictions scoring 4.1%. In the same manner, 44 cases are rightly predicted as Patients from the set of total 45 Patients predictions attaining the accuracy of 97.8%, along with 1 negative predictions of Patient as Control accounting 2.2%.

Out of 48 total Controls 47 are correctly predicted bestowing 97.9% correct predictions. With 1 Control being wrongly predicted as Patient contributing 2.1%. In the same manner, amongst 46 Patients, 95.7% are correctly predicted as Patients computing to 44 positive predictions, with 2 of the wrong predictions proportioning to 4.3%.

All inclusive, accuracy of 96.8% is achieved with 3.2% mis-classification. Explicitly, out of 96 total instances, 91 are correctly predicted by the model that makes 96.8% along with 3 incorrect predictions resulting in 3.2%.

In addition the AUC in the ROC-curve computes to 0.97985 depicting the adequacy of the model for positive predictions.

Besides accuracy 3 other evaluation metrices are considered including Precision of 96.85% with 96.78% Recall and finally the F1-Score of 0.9682 is achieved. These scores demonstrate the efficacy of the model. The values of Precision and Recall are almost the same which shows the effectiveness of the predictions made by the model. With F1-Score demonstrating the relativity of precision and recall which is recorded with high value among all iterations.

vi.) On-med Patients vs Healthy Control

In the UCSD dataset, the results are far better than the UNM dataset as the training and the confusion matrix describe the validation accuracy of 98.95% in accordance with 97.90% of testing accuracy shown by the model. This is better than as achieved regarding the off-med patients and in comparison, to all the proposed models including the Precision of 98%, Recall of 97.87%, and the F1-score of 0.9794. the results are shown in the figures Fig 11a, Fig 11b, Fig 11c, and 11d.

The first case we have considered in the UCSD dataset is On-med. In a similar fashion, the confusion matrix obtained in this case too describes the model performance precisely. The diagonal elements state the correct predictions made by the model in the form of True Positives (tp) and True Negatives (tn). In the On-med circumstance, there are 48 cases that are correctly predicted as Controls from the total 95 cases yielding 50.5%. Also, 45 cases are correctly predicted as Patients contributing 47.4% among the entire collection.

In the same manner there are 2 cases where the model miss-classifies Patients as Controls rendering 2.1% from the whole set of readings. On the other hand, there are not any incorrect predictions for the controls as Patients. 48 of the total Control predictions are correctly predicted as Controls notably 96% of correct predictions. On the same side, 2 of the total predictions are inaccurate that is 4% of the cases are incorrectly predicted. Together with 45, correct predictions are made among the total of 45 total patients cases accomplishing 100% predictions with not any wrong predictions.

Out of the 48 total Controls all the 48 are correctly predicted as Controls which means 100% correct predictions without any wrong prediction. Similarly, out

of 47 Patients, 45 are correctly predicted as patients contributing 95.7% with 2 incorrect predictions as Controls resulting in 4.3% incorrect predictions.

AUC in the ROC-curve calculates to 0.98958 depicting the dominance of positive predictions over wrong predictions.

On the whole, 97.9% are the correct predictions whereas 2.1% are inaccurate predictions achieving in 97.9% of accuracy. In addition, the evaluation metrics were achieved in the form of Precision of 98%, Recall of 97.87% along with F1-Score of 0.9794 and the AUC 0.98958. The evaluation metrics with high values of Precision, Recall, and the F1-Score describe the efficiency of the proposed model.

Across the board all the results are discussed and presented in detail. It could be observed the model doesn't respond well in the case of the UNM dataset. Achieving only maximum accuracy of 94.90% for validation checks and the testing accuracy of 93.10% amongst all the cases within the UNM dataset specifically for the Off-med Eyes Open instance. For the UCSD dataset, we get maximum accuracy in the case of on-med patients with validation accuracy of 98.5% with testing accuracy of 97.9%.

V. CONCLUSION

We have proposed a methodology consisting of 23 layered CNN for classifying PD from the healthy controls. This includes six variants from the two datasets. In all those cases the proposed methodology has proven better than state-of-art techniques as demonstrated in Tables 3 and 4 while having equal performance metrics in some cases. Our main concern was to distinguish between the PD patients and the Healthy controls in all the cases exclusively, achieving better performance metrics. Although, in most of the discussed approaches none of them have taken all the respective cases, especially in the scenario of the UNM dataset which is composed of four cases shown in Table 2. In all these cases the proposed methodology had achieved an accuracy of above 90%, and other performance evaluation matrices also concluded with a score of more than 90%. On the other hand, in the case of the UCSD dataset, we have attained performance evaluation metrics up to 98% which is better than recent research.

REFERENCES

[1] G. E. Alexander, "Biology of Parkinson's disease: Pathogenesis and pathophysiology of a multisystem neurodegenerative disorder," *Dialogues Clin. Neurosci.*, vol. 6, no. 3, pp. 259–280, Sep. 2004.

[2] S. Q. A. Rizvi, G. Wang, and X. Xing, "Early detection of Parkinson disease using wavelet transform along with Fourier transform," in *Proc. Int. Conf. Smart City Informatization*. Cham, Switzerland: Springer, 2019, pp. 323–333.

[3] T. Sood and P. Khandnor, "Classification of Parkinson's disease using various machine learning techniques," in *Int. Conf. Adv. Comput. Data Sci.* Cham, Switzerland: Springer, 2019, pp. 296–311.

[4] H. C. Verma, T. Ahmed, S. Rajan, M. K. Hasan, A. Khan, H. Gohel, and A. Adam, "Development of LR-PCA based fusion approach to detect the changes in mango fruit crop by using Landsat 8 OLI images," *IEEE Access*, vol. 10, pp. 85764–85776, 2022.

[5] R. He, X. Yan, J. Guo, Q. Xu, B. Tang, and Q. Sun, "Recent advances in biomarkers for Parkinson's disease," *Frontiers Aging Neurosci.*, vol. 10, p. 305, Oct. 2018.

[6] M. K. Hasan, T. M. Ghazal, A. Alkhalifah, K. A. Abu Bakar, A. Omidvar, N. S. Nafi, and J. I. Agbinya, "Fischer linear discrimination and quadratic discrimination analysis–based data mining technique for Internet of Things framework for healthcare," *Frontiers Public Health*, vol. 9, Oct. 2021, Art. no. 737149.

[7] M. K. Hasan, S. Islam, I. Memon, A. F. Ismail, S. Abdullah, A. K. Budati, and N. S. Nafi, "A novel resource oriented DMA framework for Internet of Medical Things devices in 5G network," *IEEE Trans. Ind. Informat.*, vol. 18, no. 12, pp. 8895–8904, Dec. 2022.

[8] A. M. Maitín, A. J. García-Tejedor, and J. P. R. Muñoz, "Machine learning approaches for detecting parkinson's disease from eeg analysis: A systematic review," *Appl. Sci.*, vol. 10, no. 23, p. 8662, 2020.

[9] S. Sivaranjini and C. Sujatha, "Deep learning based diagnosis of parkinson's disease using convolutional neural network," *Multimedia Tools Appl.*, vol. 79, no. 21, pp. 15467–15479, 2020.

[10] A. Khan, J. P. Li, M. K. Hasan, N. Varish, Z. Mansor, S. Islam, R. A. Saeed, M. Alshammari, and H. Alhumyani, "PackerRobo: Model-based robot vision self supervised learning in CART," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 12549–12566, Dec. 2022.

[11] W. Khan, M. Haroon, A. N. Khan, M. K. Hasan, A. Khan, U. A. Mokhtar, and S. Islam, "DVAEGMM: Dual variational autoencoder with Gaussian mixture model for anomaly detection on attributed networks," *IEEE Access*, vol. 10, pp. 91160–91176, 2022.

[12] L. Hu and Z. Zhang, *EEG Signal Processing and Feature Extraction*. Cham, Switzerland: Springer, 2019.

[13] A. Vavoulis, P. Figueiredo, and A. Vourvopoulos, "A review of online classification performance in motor imagery-based brain–computer interfaces for stroke neurorehabilitation," *Signals*, vol. 4, no. 1, pp. 73–86, 2023.

[14] K. Abbas, M. K. Hasan, A. Abbasi, S. Dong, T. M. Ghazal, S. N. H. S. Abdullah, A. Khan, D. Alboaneen, F. R. A. Ahmed, T. E. Ahmed, and S. Islam, "Co-evolving popularity prediction in temporal bipartite networks: A heuristics based model," *IEEE Access*, vol. 11, pp. 37546–37559, 2023.

[15] C. Dhasarathan, M. K. Hasan, S. Islam, S. Abdullah, U. A. Mokhtar, A. R. Javed, and S. Goundar, "COVID-19 health data analysis and personal data preserving: A homomorphic privacy enforcement approach," *Comput. Commun.*, vol. 199, pp. 87–97, Feb. 2023.

[16] S. Verma, T. Goel, M. Tanveer, W. Ding, R. Sharma, and R. Murugan, "Machine learning techniques for the schizophrenia diagnosis: A comprehensive review and future research directions," 2023, *arXiv:2301.07496*.

[17] G. Mirzaei and H. Adeli, "Machine learning techniques for diagnosis of Alzheimer disease, mild cognitive disorder, and other types of dementia," *Biomed. Signal Process. Control*, vol. 72, Feb. 2022, Art. no. 103293.

[18] A. Khosla, P. Khandnor, and T. Chand, "Automated diagnosis of depression from EEG signals using traditional and deep learning approaches: A comparative analysis," *Biocybern. Biomed. Eng.*, vol. 42, no. 1, pp. 108–142, Jan. 2022.

[19] O. K. Cura and A. Akan, "Classification of epileptic EEG signals using synchrosqueezing transform and machine learning," *Int. J. Neural Syst.*, vol. 31, no. 5, May 2021, Art. no. 2150005.

[20] M. B. Qureshi, M. Afzaal, M. S. Qureshi, and M. Fayaz, "Machine learning-based EEG signals classification model for epileptic seizure detection," *Multimedia Tools Appl.*, vol. 80, no. 12, pp. 17849–17877, May 2021.

[21] Y.-G. Hong, H.-K. Kim, Y.-D. Son, and C.-K. Kang, "Identification of breathing patterns through EEG signal analysis using machine learning," *Brain Sci.*, vol. 11, no. 3, p. 293, Feb. 2021.

[22] M. K. M. Rabby, A. K. M. K. Islam, S. Belkasim, and M. U. Bikdash, "Epileptic seizures classification in EEG using PCA based genetic algorithm through machine learning," in *Proc. ACM Southeast Conf.*, Apr. 2021, pp. 17–24.

[23] N. Betrouni, A. Delval, L. Chaton, L. Defebvre, A. Duits, A. Moonen, A. F. G. Leentjens, and K. Dujardin, "Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results," *Movement Disorders*, vol. 34, no. 2, pp. 210–217, Feb. 2019.

[24] M. Chaturvedi, F. Hatz, U. Gschwandtner, J. G. Bogaarts, A. Meyer, P. Fuhr, and V. Roth, "Quantitative EEG (QEEG) measures differentiate Parkinson's disease (PD) patients from healthy controls (HC)," *Frontiers Aging Neurosci.*, vol. 9, p. 3, Jan. 2017.

[25] E. Arasteh, A. Mahdizadeh, M. Mirian, S. Lee, and M. McKeown, "Deep transfer learning for Parkinson's disease monitoring by image-based representation of resting-state EEG using directional connectivity," *Algorithms*, vol. 15, no. 1, p. 5, Dec. 2021.

[26] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Comput. Appl.*, vol. 32, no. 24, pp. 10927–10933, 2020.

[27] M. Murugappan, W. Alshuaib, A. K. Bourisly, S. K. Khare, S. Sruthi, and V. Bajaj, "Tunable Q wavelet transform based emotion classification in Parkinson's disease using electroencephalography," *PLoS ONE*, vol. 15, no. 11, Nov. 2020, Art. no. e0242014.

[28] M. Bugdol, D. Ledwoń, M. N. Bugdol, K. Zawiślak-Fornagiel, M. Danch-Wierzchowska, and A. W. Mitas, "Eeg signal and deep learning approach in evaluation of cognitive declines in parkinson's disease," in *Proc. Int. Conf. Inf. Technol. Biomed.* Cham, Switzerland: Springer, 2022, pp. 43–53.

[29] S. Hassin-Baer, O. S. Cohen, S. Israeli-Korn, G. Yahalom, S. Benizri, D. Sand, G. Issachar, A. B. Geva, R. Shani-Hershkovich, and Z. Peremen, "Identification of an early-stage Parkinson's disease neuromarker using event-related potentials, brain network analytics and machine-learning," *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0261947.

[30] R. Zhang, J. Jia, and R. Zhang, "EEG analysis of Parkinson's disease using time–frequency analysis and deep learning," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103883.

[31] R. Sugden and P. Diamandis, "Generalizable electroencephalographic classification of parkinson's disease using deep learning," *MedRxiv*, vol. 42, Sep. 2023.

[32] M. Shaban and A. W. Amara, "Resting-state electroencephalography based deep-learning for the detection of Parkinson's disease," *PLoS ONE*, vol. 17, no. 2, Feb. 2022, Art. no. e0263159.

[33] C. Chu, Z. Zhang, Z. Song, Z. Xu, J. Wang, F. Wang, W. Liu, L. Lu, C. Liu, X. Zhu, C. Fietkiewicz, and K. A. Loparo, "An enhanced EEG microstate recognition framework based on deep neural networks: An application to Parkinson's disease," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 3, pp. 1307–1318, Mar. 2023.

[34] J. F. Cavanagh, A. Napolitano, C. Wu, and A. Mueen, "The patient repository for EEG data + computational tools (PRED+CT)," *Frontiers Neuroinform.*, vol. 11, p. 67, Nov. 2017.

[35] A. P. Rockhill, N. Jackson, J. George, A. Aron, and N. C. Swann, "UC San Diego resting state EEG data from patients with Parkinson's disease," Univ. Oregon, Eugene, OR, USA, Tech. Rep., 2020, doi: 10.18112/openneuro.ds002778.v1.0.2.

[36] S. J. Luck, *An Introduction to the Event-Related Potential Technique*. Cambridge, MA, USA: MIT Press, 2014.

[37] M. Z. I. Ahmed, N. Sinha, E. Ghaderpour, S. Phadikar, and R. Ghosh, "A novel baseline removal paradigm for subject-independent features in emotion classification using EEG," *Bioengineering*, vol. 10, no. 1, p. 54, Jan. 2023.

[38] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, p. 987, Feb. 2019.

[39] R. Ghosh, S. Phadikar, N. Deb, N. Sinha, P. Das, and E. Ghaderpour, "Automatic eyeblink and muscular artifact detection and removal from EEG signals using k-nearest neighbor classifier and long short-term memory networks," *IEEE Sensors J.*, vol. 23, no. 5, pp. 5422–5436, Mar. 2023.

[40] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights Into Imag.*, vol. 9, no. 4, pp. 611–629, Aug. 2018.

[41] S. K. Khare, V. Bajaj, and U. R. Acharya, "Detection of Parkinson's disease using automated tunable Q wavelet transform technique with EEG signals," *Biocybern. Biomed. Eng.*, vol. 41, no. 2, pp. 679–689, Apr. 2021.

[42] S. A. P. Ruiz, "Parkinson's disease diagnosis using electroencephalographic (EEG) signal processing and machine learning techniques," B.S. thesis, School Math. Comput. Sci., Universidad de Investigación de Tecnología Exp. Yachay, San Jose, CA, USA, 2021.

[43] L. Qiu, J. Li, and J. Pan, "Parkinson's disease detection based on multi-pattern analysis and multi-scale convolutional neural networks," *Frontiers Neurosci.*, vol. 16, Jul. 2022, Art. no. 957181.

[44] M. F. Anjum, S. Dasgupta, R. Mudumbai, A. Singh, J. F. Cavanagh, and N. S. Narayanan, "Linear predictive coding distinguishes spectral EEG features of Parkinson's disease," *Parkinsonism Rel. Disorders*, vol. 79, pp. 79–85, Oct. 2020.

[45] S.-B. Lee, Y.-J. Kim, S. Hwang, H. Son, S. K. Lee, K.-I. Park, and Y.-G. Kim, "Predicting Parkinson's disease using gradient boosting decision tree models with electroencephalography signals," *Parkinsonism Rel. Disorders*, vol. 95, pp. 77–85, Feb. 2022.

[46] D. Shah, G. K. Gopika, and N. Sinha, "Analysis of EEG for Parkinson's disease detection," in *Proc. IEEE Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2022, pp. 1–5.

SYED QASIM AFSER RIZVI received the B.Sc. degree in information technology from Aligarh Muslim University (AMU), Aligarh, India, in 2010, and the M.C.A. degree from Bangalore University, Bengaluru, India, in 2013. He is currently pursuing the Ph.D. degree with Guangzhou University, Guangzhou, China. His research interests include machine learning (ML), signal processing, biomedical signals, and deep learning.

GUOJUN WANG (Member, IEEE) received the B.Sc. degree in geophysics, the M.Sc. and Ph.D. degrees in computer science from Central South University, Changsha, China, in 1992, 1996, and 2002, respectively. He was a Professor with Central South University; a Visiting Scholar with Temple University, Philadelphia, PA, USA, and Florida Atlantic University Boca Raton, FL, USA; a Visiting Researcher with the University of Aizu, Aizuwakamatsu, Japan; and a Research Fellow with The Hong Kong Polytechnic University, Kowloon, Hong Kong. He is currently a Pearl River Scholarship Distinguished Professor with Guangzhou University, Guangzhou, China. His research interests include cloud computing, trusted computing, and information security. He is a Distinguished Member of CCF and a member of ACM and IEICE.

ASIF KHAN (Member, IEEE) received the B.Sc. (Hons.) and M.C.A. degrees from Aligarh Muslim University, India, and the Ph.D. degree (Hons.) in computer science and technology from the University of Electronic Science and Technology of China (UESTC), China, in 2016. He did Post-doctoral Scientific Research Fellow with UESTC. He was an Adjunct Faculty with the University of Bridgeport, USA, for the China Program, in Summer 2016. Previously, he was a Visiting Scholar of big data mining and application with the Chongqing Institute of Green and Intelligent Technology (CIGIT), Chinese Academy of Sciences, Chongqing, China. He is currently an Assistant Professor with Integral University, India. He is a contributor to many international journals with robotics and vision analyses about the contemporary world in his articles. His research interests include machine learning, robotics vision, and new ideas regarding vision-based information critical theoretical research. He received the Academic Achievement Award and Excellent Performance Award by UESTC, from 2015 to 2016.

MOHAMMAD KAMRUL HASAN (Senior Member, IEEE) received the Ph.D. degree in electrical and communication engineering from the Faculty of Engineering, International Islamic University, Malaysia, in 2016. He is currently a Senior Lecturer with the Faculty of Information Science and Technology, Center for Cyber Security, Universiti Kebangsaan Malaysia (UKM). He is a certified professional technologist in Malaysia. He is specialized in elements pertaining to cutting-edge information-centric networks; computer networks, data communication and security, mobile network and privacy protection, cyber-physical systems, Industrial IoT, transparent AI, and electric vehicles networks. He has published more than 200 indexed papers in ranked journals and conference proceedings. He is a Senior Member of the Institute of Electrical and Electronics Engineers and a member of the Institution of Engineering and Technology and the Internet Society. He served as the Chair for the IEEE Student Branch, from 2014 to 2016. He has actively participated in many events/workshops/training for IEEE and IEEE Humanity programs, Malaysia. He is the general chair, the co-chair, and a speaker of conferences and workshops for the shake of society and academy knowledge building and sharing and learning. He is an Editorial Member in many prestigious high-impact journals, such as IEEE, IET, Elsevier, Frontier, and MDPI. He has been contributing and working as a volunteer for underprivileged people for the welfare of society.

TAHER M. GHAZAL (Senior Member, IEEE) received the B.Sc. degree in software engineering from Al Ain University, in 2011, the M.Sc. degree in information technology management from The British University in Dubai, which is associated with The University of Manchester and The University of Edinburgh, in 2013, the Ph.D. degree in IT/software engineering from Damascus University, in 2019, and the Ph.D. degree in information science and technology from Universiti Kebangsaan Malaysia, in 2023. With over a decade of extensive and diverse experience, he was an Instructor, a Tutor, a Researcher, a Teacher, a IT Support/Specialist Engineer, and a Business/Systems Analyst. He was with various departments, including Engineering, Computer Science, and ICT, the Head of the STEM and Innovation, and has also been involved in quality assurance, accreditation, and data analysis, in several governmental and private educational institutions under KHDA, Ministry of Education, and Ministry of Higher Education and Scientific Research, United Arab Emirates. His research interests include the IoT, IT, artificial intelligence, information systems, software engineering, web development, building info, modeling, quality of education, management, big data, quality of software, and project management. He is actively involved in community services in the projects and research field.

ATTA UR REHMAN KHAN (Senior Member, IEEE) is currently an Associate Professor with the College of Engineering and Information Technology, Ajman University, United Arab Emirates. Prior to joining AU, he was a Postgraduate Program Coordinator with Sohar University Oman; the Director of the National Cybercrime Forensics Laboratory, Pakistan; the Head of the Cybersecurity Center, Air University, Pakistan; and the Research Committee Chair of King Saud University, Saudi Arabia. He has extensive experience in teaching, research, and industry at key positions. He has developed multiple degree programs, published research articles in reputed journals/conferences, and edited/coauthored multiple books. Some of his publications have appeared among the most popular articles of the IEEE ComSoc Society, and among the top 1% highly cited papers of the academic field of computer science (citations received with respect to publication year) as per Web of Science.

Moreover, he is a Program Evaluator with the Pakistan's National Computing Education Accreditation Council (NCEAC), a member of Amazon Web Services (AWS) Educate program, NVIDIA GPU Educators Program, and a steering committee member/the pc chair/track chair/technical program committee (TPC) member of over 80 internationalÂ conferences. According to a Stanford University report, he is among the World's Top 2% Scientists. He served as the Conferences Chair for the IEEE Islamabad Section, Pakistan. He is serving on the Editorial Board of *Cluster Computing* (Springer), *IEEE Communications Magazine*, *Journal of Network and Computer Applications* (Elsevier), *The Computer Journal* (Oxford), IEEE SDN Newsletter, IEEE Access, *KSII Transactions on Internet and Information Systems*, and *Ad Hoc and Sensor Wireless Networks*.