Dicle University
Journal of Engineering
https://dergipark.org.tr/tr/pub/dumf
duje.dicle.edu.tr

Research Article

Automatic Detection of Mild Cognitive Impairment from EEG Recordings Using Discrete Wavelet Transform Leader and Ensemble Learning Methods

Afrah SAID[1*], Hanife GÖKER[2]

[1] Dumlupınar University, Electrical-Electronics Engineering Department, afrah.said@ogr.dpu.edu.tr, Orcid No: 0000-0003-1016-6429.
[2] Gazi University, Health Services Vocational College, gokerhanife@gazi.edu.tr, Orcid No: 0000-0003-0396-7885.

ARTICLE INFO

Article history:

Received 31 December 2022
Received in revised form 23 February 2023
Accepted 3 March 2023
Available online 23 March 2023

Keywords:

Discrete wavelet transform leader, EEG, ensemble learning, MCI, automatic detection

Doi: 10.24012/dumf.1227520

* Corresponding author

ABSTRACT

Mild Cognitive Impairment (MCI) is a risk of cognitive decline, commonly referred to as a transitional stage between normal cognition and dementia. Patients with MCI typically progress to Alzheimer's disease (AD), which causes cognitive deficits such as deterioration of their thinking abilities. This study aims to detect MCI patients using electroencephalography (EEG) signals. The EEG dataset used in this study consists of EEG signals recorded from 18 MCI and 16 control groups. Firstly, EEG signals were denoised using multiscale principal component analysis (multiscale PCA). Then, 36 features were extracted from the EEG signals using the discrete wavelet transform leader (DWT leader) feature extraction method. Finally, using the extracted feature vectors, control groups, and MCI groups were classified by ensemble learning algorithms. As a result, AdaBoostM1 algorithm gained the highest performance with 93.50% accuracy, 93.27% sensitivity, 93.75% specificity, 94.38% precision, 93.82% f1-score, and 86.97% Matthews correlation coefficient (MCC). By achieving quite satisfactory accuracy, this study proves that the ensemble learning algorithm can also be used for MCI detection.

Introduction

MCI is a stage between normal brain aging and dementia. It is a condition in which patients have a decline in short-term memory, along with forgetfulness about recent events. However, without obvious impairment in everyday functioning [1-2]. Age-related diseases are becoming more common around the world as a result of the progressive increase in lifespans. The most popular among age-related diseases are the MCI and AD. MCI is the early stage of dementia that could also be referred to the “mild cognitive disorder” classification category under the International Classification of Diseases by the World Health Organization [3-4].

Patients with MCI typically develop AD and those patients normally experience problems with motor, cognitive, and behavioral symptoms. Furthermore, they face difficulties in adapting to changes and external stressors. Consequently, this could lead them to a need for a caregiver due to their behavioral disturbances [5]. Therefore, MCI is a crucial step for the early identification of AD. This is because patients with MCI are more likely to develop AD than those without MCI. For that reason, early diagnosis will help patient selection for upcoming clinical trials. It will also motivate the patients to implement a new lifestyle. Furthermore, it can assist the patient’s family and the patients themselves in considering financial planning for future care needs. This is because MCI rates are increasing rapidly with 5% of the general population and around 15% developing dementia each year [3, 6-7].

For the detection of MCI, patients are generally checked in medical care using magnetic resonance imaging. This method is cost-effective and less effective in the early detection of MCI as its performance is limited to the late stage of dementia. Another technique used is positron emission tomography which is the most popular neuroimaging technique for diagnosing dementia. This method can show the presence of the amyloid protein, a protein in the brain that is associated with dementia, and neurometabolic abnormalities in the brain. However, this diagnosis requires special equipment which is expensive to perform. Moreover, it involves exposure to radiation which is harmful [2]. To avoid the expensive common diagnosis techniques mentioned above, recently, researchers are developing predictive modeling techniques. This is done by using the information of the brain activity of the patient, collected from EEG devices. Moreover, machine learning methods are also used due to their efficiency and low cost for the detection of MCI and AD.

Various investigations have been carried out for the early diagnosis of MCI. Kashefpoor et.al (2019) developed a new supervised dictionary-learning model called Correlation-based Label Consistent. This method is based on the analysis of the EEG signals using k-means and the singular value decomposition method. First, they extracted spectral features from the EEG signals applied frequency and time domains. After that, they started voting between the labels to have a final label for all the channels of each EEG signal. As a result, they achieved an accuracy of 88.9% [8]. Alvi et.al (2022) proposed a Long Short-Term Memory (LSTM) based framework for the early detection of MCI from EEG signals. To conclude their model performance, they designed 20 different LSTM models with 5-fold cross-validation and as a result, they achieved an accuracy of 91.41% [9]. Jamaloo et.al (2020) investigated the frequency bands of the EEG signals to distinguish between MCI and healthy subjects using hidden markov model. A model with hidden states. Firstly, they divided the EEG signals into 5 based on standard frequency bands, the delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–25 Hz), and gamma (25–35 Hz). After that, using leave one out cross-validation method, they divided them into train, test, and validation sets. As a result, they have seen and concluded that the alpha and gamma frequency bands produced the highest classification accuracy with %95.9 ± 0.4 alpha and %97.2 ± 0.5 Gamma [10].

Different from the above studies, a new model was proposed using ensemble learning algorithms that detect MCI patients using EEG signals. The accomplishments of this article are listed below.

- MCI and healthy subjects were classified with ensemble learning algorithms using feature vectors of the EEG signals that were extracted using the DWT leader feature extraction method.
- The best performance model was chosen after evaluating all the ensemble learning algorithms models.
- As a result, the proposed model was able to enhance classification accuracy.

Material and Method

Proposed Method

This study is carried out by first capturing the EEG signals of the subjects using an EEG cap. Then the EEG signals of each of the subjects were segmented into three parts of 10 minutes each. After the segmentation process, the EEG signals were explored using Multiscale PCA. This is because it looks for the subspaces that maximize the sum of all squared pairwise distances between data projections [11]. Then DWT leader feature extraction method has been used to extract the feature vectors of the EEG signals. Subsequently, 36 features were extracted using the DWT leader.

After that, ensemble learning algorithms such as bagging, AdaboostM1, Gentle Boost, Logit Boost, LP Boost, Robust Boost, RUS Boost, and Total Boost ensemble learning algorithms were used. These algorithms have been used to classify the subjects using the spectral feature vectors that have been extracted in the feature extraction step. After classifying them, the performances of all the algorithms are analyzed to examine the algorithm with the highest performance. As well as to reach a conclusion of which algorithm was efficient in detecting MCI. Moreover, the implementation steps of the proposed model are shown in Figure 1.

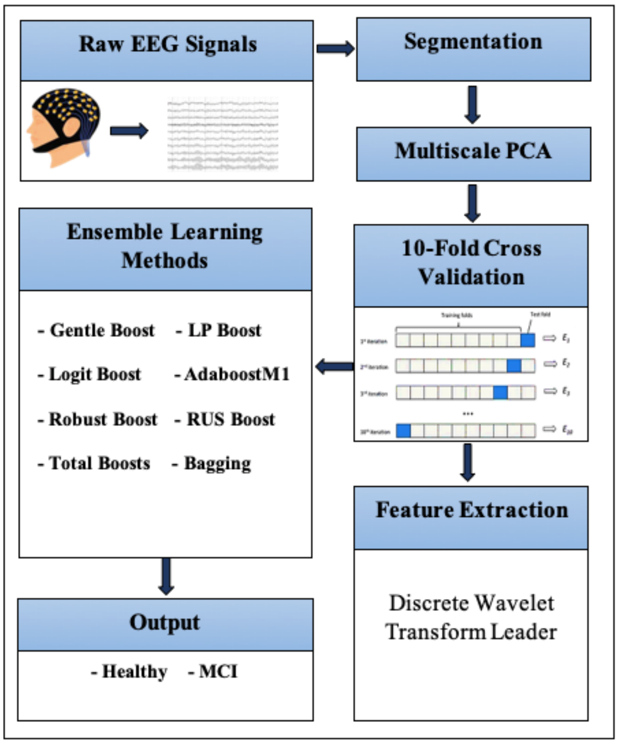

Figure 1: The implementation steps of the proposed model.

Dataset

To conduct this study, a dataset that is publicly available on the internet that consists of EEG signals from 34 subjects (16 control and 18 MCI) has been used [8]. Where each participant had at least completed their upper or elementary education, and their ages ranged from 40 to 77. Each participant's EEG was continuously recorded for 30 minutes with a skin-electrode impedance of less than 5k and a 256 sampling rate. In addition to 19 electrodes positioned according to the 10-20 International System (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2). Moreover, Mini-Mental State Examination (MMSE) tests were also conducted. Considering the score of MMSE between 21 and 26 as an MCI patient and scores greater than that being healthy. Additionally, the study process of the EEG signals has gained ethical approval from Isfahan University of Medical Sciences' deputy for research and technology. Moreover, written consent was taken from the participants concerning the description of the study procedure [8]. Figure 2 depicts electrode positioning for the international 10-20 system.

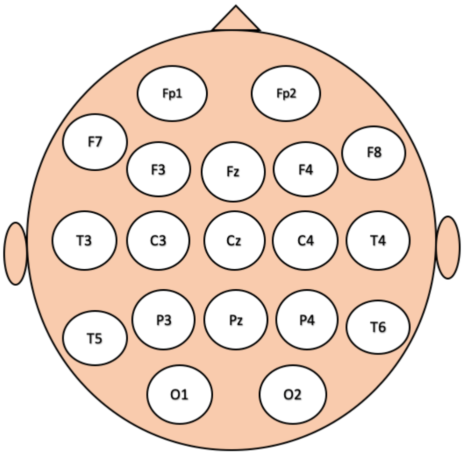

Figure 2: International 10-20 system electrode placement system.

Segmentation Process and Multiscale PCA

In the segmentation process, each of the 34 participants of the EEG recordings consists of 30 minutes with 256 sampling rates. Meaning that there are 19 electrodes with 460800 (256 x 60 x 30) rows for each participant. Therefore, processing this enormous dataset requires computational support and requires an expense of time. For that reason, we have segmented each of the recordings into three segments where each segment consisting of 10 minutes. As a result, there are 3 segments for each subject and a total of 102 segments is produced with a size of 153600 (256 Hz x 60 seconds x 10 minutes) samples x 19 channels. Correspondingly the total number of the dataset is 1938 (3 segments x 34 subjects x 19 channels). The segmentation process was carried out to increase the number of dataset data so that the developed dataset can be practical and computationally efficient. Moreover, the dataset is segmented to restore the spectral and temporal characteristics of EEG. Furthermore, to prevent data loss when feature extraction is performed on the EEG signal which has a very large length [12-13]. EEG signals are nonstationary and this means that they are vulnerable to noises. Therefore, there must be a method to eliminate or reduce the noise from the signals. This can be done using de-noising methods such as band-pass filtering and PCA. However, the use of the filtering technique depends on the nature of the signals. Accordingly, in this study as the signals are based on wavelets, multiscale PCA was used. After segmenting the dataset, the multiscale PCA was used which is the combination of potential PCA and orthonormal wavelets. The orthonormal wavelets are responsible for splitting the stochastic processes and the PCA for extracting the relationship between multiple variables [12].

K-Fold Cross Validation

For the division of the dataset into training and test sets, the k-fold cross-validation method has been used. K-fold cross-validation is a technique to evaluate the results produced by the model by dividing the dataset into a training set to train the model and a testing set to evaluate it [14]. It is used to achieve an unbiased estimate of the model performance by dividing the dataset into k subsets of equal size. Therefore, building models k times and every time removing out one of the subsets from the training set and then using it as the test set. It is also used to minimize errors in the model [15]. The characteristics of the dataset are shown in Table 1.

Table 1. The characteristics of the dataset.

Class	Number of subjects	Total number of datasets
MCI	18	1026
Healthy	16	912
Total	34 subjects x 19 channels x 3 segments =1938	

Discrete Wavelet Transform Leader (DWT Leader) Feature Extraction

Feature extraction methods are used to minimize irrelevant features. Thereby only extracting the important features from the EEG signals to increase the performance of the classification model and avoid overfitting of the data which could happen without feature extraction [16-17]. EEG signals are non-stationary and obtaining frequency information during brain activity is a challenging process. This is due to electromagnetic interference between the oscillators' high frequency and the low frequency produced by eye blinks. Thus, classifying the raw signals eventually affects the accuracy of the model [16]. Therefore, in this study DWT leader feature extraction method was used to extract the most important features and to estimate the singularity of the spectrum. Wavelet transform is used for time-frequency analysis. Whereas, DWT is a feature extraction method that analyzes the signal with different resolutions by breaking down the approximation and detailed coefficient signals in different frequency bands. When resolving high-frequency components in a minor window, the DWT method requires large time windows to solve low-frequency components. Because the signal consists of both low and high frequency components it attempts to provide the best resolution in terms of both time and frequency [18]. Recently, a new form called DWT leader was built from DWT by Jaffard and his co-workers. DWT leader is a method that works well with non-stationary and non-linear signals [19]. DWT leader is the local suprema of the wavelet coefficient of the signal. This method was introduced to improve the usual wavelet methods. This method uses wavelet analysis to estimate the multifractal spectrum which is a tool for the analysis of fluctuations. Given that it describes the unique behavior of a signal. The DWT is an effective time-frequency analysis tool and its wavelet coefficients are suitable measurements to examine its regularity. However, its drawback is that it fails for signals with fluctuating singularities. The wavelet leader method fixed that drawback. Moreover, DWT leader has low computational complexity and offers robust and fast estimation. Furthermore, it performs better and is more effective than DWT at capturing oscillating signals [20-23]. The DWT of a signal X(t) is shown in equation 1, whereas the local supremum of the wavelet coefficients in the dyadic cube is known as the DWT leader and it is shown in equation 2 [19].

$$d_X(j,k) = \int_R X(t)2^{-j}\psi(2^{-j}t-k)dt \tag{1}$$

$$L_X(j,k) = sup_{\lambda' \sqsubset \Gamma}|d_{X,\lambda}| \tag{2}$$

Whereas, λ is the dyadic interval and the Γ is the dyadic cube and their equations are as follows [19].

$$\lambda = [k.2^j,(k+1)2^j] \tag{3}$$

$$\Gamma = 3\lambda = \lambda_{jk-1} \cup \lambda_{jk} \cup \lambda_{jk+1} \tag{4}$$

In this study, using the DWT leader as a feature extraction method, 36 features were extracted from the EEG signals.

Ensemble Learning Classification Algorithms

Ensemble learning is referred to as multiple classifier systems where multiple learners are trained and combined to address a learning problem [24]. Ensemble learning helps improve model performance. It is the most well-known and comprehensive machine learning field. This is due to its better performance than machine learning algorithms [25]. The framework of the ensemble learning method works by first training a group of learners individually. Then merging them through some strategies thereby gaining high performance. Whereas the reason behind its high performance is that, the generalization ability of an ensemble group of learners is stronger than that of an individual learner [24].

Boosting

Boosting algorithms begin their process by first training a base learner and then modifying the distribution of the training samples. This depends on the base learner's performance so that samples that were mistakenly classified would be given more focus by subsequent base learners. The training sample distribution is again modified once the second base learner has been trained using the first base learner's adjusted training samples. Moreover, the process is repeated again and again until a certain number of base learner's "T" has been reached. As a result, the base learners are combined and weighted [24]. AdaBoost is the most popular boosting algorithm where the additive model boosting equation is shown in equation (5).

$$H(x) = \sum_{t=1}^{T} \alpha_t h_t(x) \tag{5}$$

Where α_t is the target t, h_t (x) is the chosen base learner at each stage to minimize the loss function [24].

Bagging

Bagging is an effective and useful ensemble learning algorithm used in classification and regression models [26]. It generates sampling subsets by randomly picking and copying one sample from the original dataset to the sampling subset. Then the process is repeated several times. Then using those subsets, the base classifiers are trained and combined, where the training of the basic algorithm is performed in a parallel way [24, 26]. The combining method uses the simple voting approach for classification problems. While for regression tasks, it uses the simple averaging method. The equation of the bagging algorithm is shown in equation (6).

$$H^{oob}(x) = \underset{\gamma \in \Upsilon}{arg}\, arg\, max \sum_{t=1}^{T} |(h_t(x) = y). \tag{6}$$

$$|(x \notin D_t)$$

Where, D_t represents the set of samples used by the learner h_t, and $H^{oob}(x)$ represents the prediction of sample x of out-of-bag. Thereby, only taking into consideration of the predictions made by base classifiers that did not use the sample x for training [24].

Performance Evaluation Metrics

The efficiency of a model that is proposed for the classification depends on counting all the correct predictions that are made from all the predictions that were made [27]. Therefore, to assess how well the classification model performed at predicting whether the subject had MCI or not, performance evaluation metrics using confusion matrix parameters were used and their equations are shown below. Where TP, FN, FP, and TN stand for True Positive, False Negative, False Positive, and True Negative respectively. Moreover, the values of TP and TN represent how many samples were correctly identified whereas the values of FP and FN represent how many samples were misclassified.

Precision (Prec)=TP/(TP+FP) (7)

Accuracy (ACC)=(TP+TN)/(TP+FP+FN+TN) (8)

Sensitivity (Sens)=TP rate =TP/(TP+FN) (9)

Specificity (Spec)=TN rate= TN/(TN+FP) (10)

F1-score=2TP/(2 TP+FP+FN) (11)

$$MCC=\frac{(TN \times TP - FP \times FN)}{\sqrt{(TN+FN)x(FP+TP)x(TN+FP)x(FN+TP)}} \tag{12}$$

Experimental Results and Discussion

After segmenting the dataset of the 34 participants into 3 segments, a total of 1938 (3 segments x 34 subjects x 19 channels) datasets were generated. Then using multiscale PCA the signals were explored and important features were extracted from the signals by the DWT leader. Following the next step is the dataset division. The dataset was divided using the k-fold cross-validation method. Where k=10, meaning one-tenth for testing and the rest for training the model. Thus, repeating the process ten times, each time using a different tenth for testing. After dividing the dataset, the model was trained with the training set and tested with the testing set. Then ensemble learning algorithms were used to classify the MCI subjects from the healthy subjects.

To check whether the classification of the subjects was accurately classified or not, confusion matrix parameters of ensemble learning algorithms such as bagging, AdaboostM1, Gentle Boost, Logit Boost, LP Boost, Robust Boost, RUS Boost, and Total Boost were calculated and the outcome of each algorithm is shown in Table2.

Table 2. Confusion matrix parameters of Ensemble Learning methods.

Ensemble Learning Methods	TN	FN	FP	TP	TP+TN	FP+FN
Bagging	857	207	55	819	1676	262
AdaboostM1	855	69	57	957	1812	126
Gentle Boost	841	191	71	835	1676	262
Logit Boost	791	154	121	872	1663	275
LP Boost	841	305	71	721	1562	376
Robust Boost	797	193	115	833	1630	308
RUS Boost	791	184	121	842	1633	305
Total Boost	752	153	160	873	1625	313

After analyzing the confusion matrix parameters of the ensemble learning algorithms that are shown in Table 2, it is seen that the AdaboostM1 algorithm has the highest proportion of correctly classified samples with 1812 and 126 inaccurately classified samples. Moreover, to evaluate the performance of the algorithms they were assessed with the evaluation performance metrics namely, sensitivity, specificity, precision, f1-score, MCC, and accuracy, and their values are shown in Table 3.

Table 3. Performance evaluation metrics of Ensemble Learning methods.

Ensemble Learning Methods	Sens.	Spec.	Prec.	F1-score	MCC	ACC
Bagging	79.82	93.97	93.71	86.21	74.02	86.48
AdaboostM1	93.27	93.75	94.38	93.82	86.97	93.50
Gentle Boost	81.38	92.21	92.16	86.44	73.63	86.48
Logit Boost	84.99	86.73	87.81	86.38	71.62	85.81
LP Boost	70.27	92.21	91.04	79.32	63.45	80.60
Robust Boost	81.19	87.39	87.87	84.40	68.48	84.11
RUS Boost	82.07	86.73	87.44	84.67	68.68	84.26
Total Boost	85.09	82.46	84.51	84.80	67.57	83.85

Table 3 was examined to check the best performance among the ensemble learning algorithms based on their accuracy value. After examining their accuracy, it is seen that the AdaboostM1 algorithm had the highest accuracy with 93.50%. Stating that the AdaboostM1 algorithm was the best algorithm for classifying MCI from healthy ones among the other algorithms. In Figure 3, the accuracy values of ensemble learning methods are given.

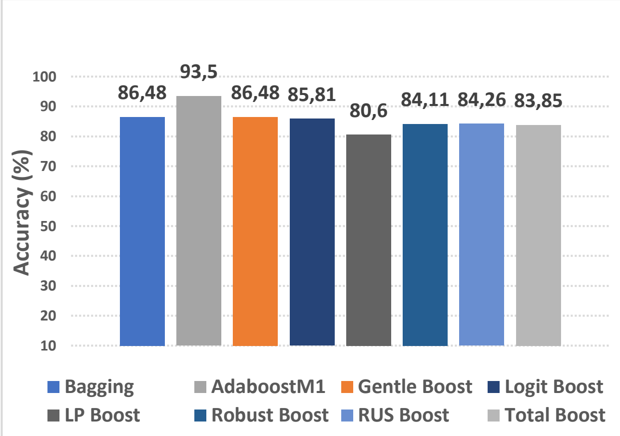

Figure 3: Accuracy performance of the classification algorithms.

AdaboostM1 algorithm had the highest accuracy (93.50%) among the other algorithms. The reason why AdaboostM1 gained the highest accuracy among the rest of the algorithms is due to its higher prediction accuracy. The AdaboostM1 algorithm is a powered variant of the Adaboost algorithm. AdaboostM1 improves the performance of classifiers for less classification errors. Also, the advantages of AdaboostM1 over other ensemble learning algorithms are its strong theoretical foundation, simplicity in implementation, high prediction accuracy, and overfit protection by training on small subdivisions of training data and weighted training data. In addition, AdaboostM1 employs an exponential loss function and provides the ability to identify complex composite classifiers from small amounts of data [28-29]. Given the benefits listed above regarding the AdaBoostM1 algorithm, it is preserved to be the best classifying algorithm for the detection of MCI. Moreover, previous studies based on MCI detection using EEG signals were compared with this study and this is shown in Table 4.

Table 4. A comparative analysis with relevant literature studies.

Study	Signal Processing	Subjects	Best Classifier	ACC
Ruiz-Gómez et al. [30]	Spectral and non-linear analyses	37 AD, 37 MCI, 37 healthy	MLP	78.43%
Poza et.al [31]	RP and SF	19 MCI, 37 AD, 29 healthy	LDA	79.2%
Hadiyoso & Tati [32]	Hjorth Descriptor	5 healthy 5 MCI	KNN	80.00%
Hadiyoso et al. [33]	Linear QEEG based power spectral features	16 healthy 11 MCI	KNN	81.5%
Kashefpoor et al. [34]	Spectral features	16 healthy 11 MCI	NF with KNN	88.89%
Proposed model	DWT leader	16 healthy 18 MCI	Adaboost M1	93.50%

RP: Relative power; SF: Spectral flux; NF: Neuro fuzzy; MLP: Multi-layer perceptron; KNN: K-nearest neighbors.

Several studies explored the detection of MCI using different methods, five of them are mentioned in this study for comparison. The study of Ruiz-Gómez et.al proposed a model using three different classification models with spectral analyses and non-linear analyses for the extraction of spectral features from the EEG signals. As a result, they have maintained to achieve an accuracy of 78.43% with the MLP model. The study of Poza et.al developed a model using LDA with RP to analyze the conventional EEG frequency bands and SF to explore the Spatio-temporal fluctuations. As a result of their research, they attained 79.2% accuracy. The studies of [33] and [34] used the same dataset that is used in this study. However, the study of [33] used a different technique for the signal processing of the EEG signals by using linear QEEG based power spectral features. However, the study of [32] used small dataset that consist of only 10 subjects, with 17 subjects less than that of study [33], [34], and 24 subjects less than that of this study. Besides that, they used a hjorth descriptor for their signal processing and they also used the same classifier that used study [33], the KNN classifier. However, the study of [33] gained a higher accuracy than the study of [32] with an accuracy of 81.5%, 1.5% greater than that of [32]. While the study of [34] used spectral features for signal processing as in this study. However, they have extracted 19 spectral features from each of the 19 channels for each of the participants. After that, they classified them using the NF system with the KNN classifier, thereby achieving an accuracy of 88.89%. The study of [30] and [31] used a different dataset than the dataset used in this study and study [32], [33], and [34]. Moreover, the number of their subjects is quite higher than the number of the subjects of this study and the study of [32], [33], and [34]. With the comparison of previous studies using the same dataset and using a different dataset, it is clear that this study outperformed all these works and maintained to achieve an accuracy of 93.50%. The cause behind the high accuracy is the employment of ensemble learning algorithms. This is because ensemble learning algorithms improve model performance by independently training a group classifier. After that, it combines them using specific techniques, which is why it outperforms machine learning algorithms.

CONCLUSION

To conclude, an ensemble learning-based model has been proposed with DWT leader as a feature extraction method to extract the feature vectors of the EEG signals. After that, we classified them using ensemble learning algorithms, namely Bagging, AdaboostM1, Gentle Boost, Logit Boost, LP Boost, Robust Boost, RUS Boost, and Total Boost. The confusion matrix parameters of the ensemble learning classification algorithms were calculated. The performance results of the ensemble learning algorithms were compared with performance evaluation metrics. As a result of the comparison, it is noticed that the AdaboostM1 algorithm gained the highest accuracy among them. It has achieved an accuracy of 93.50%, 93.27% sensitivity, 93.75% specificity, 94.38% precision, 93.82% f1-score, and 86.97% MCC. Using ensemble learning algorithms with DWT leader, we have developed a new model that can be used for the detection of MCI. This model also demonstrates that the ensemble learning algorithm can be used to detect MCI and gain quite a high accuracy. Moreover, this finding suggests that an ensemble learning algorithm-based model can help physicians to differentiate MCI from healthy groups.

Ethics committee approval

There is no need to obtain permission from the ethics committee for the article prepared.

Conflict of Interest Statement

There is no conflict of interest with any person / institution in the article prepared.

References

[1] Y. Tao, Y. Han, L. Yu, Q. Wang, S.X. Leng, and H. Zhang, “The predicted key molecules, functions, and pathways that bridge mild cognitive impairment (MCI) and Alzheimer's disease (AD),” *Frontiers in Neurology*, vol. 11, p. 233, 2020.

[2] S. J. Lim, Z. Lee, L. N. Kwon, and H. W. Chun, “Medical health records-based Mild Cognitive Impairment (MCI) prediction for effective dementia care,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 17, p. 9223, 2021.

[3] M. N. Sabbagh, M. Boada, S. Borson, M. Chilukuri, P. M. Doraiswamy, B. Dubois, and H. Hampel, “Rationale for early diagnosis of mild cognitive impairment (MCI) supported by emerging digital technologies,” *The Journal of Prevention of Alzheimer's Disease*, vol. 7, no.3, pp. 158-164, 2020.

[4] N. T. Lautenschlager, K. L. Cox, and K. A. Ellis, “Physical activity for cognitive health: what advice can we give to older adults with subjective cognitive decline and mild cognitive impairment?” *Dialogues in Clinical Neuroscience*, 2022.

[5] R. Baschi, A. Luca, A. Nicoletti, M. Caccamo, C. E. Cicero, C. D'Agate, and R. Monastero, “Changes in motor, cognitive, and behavioral symptoms in Parkinson's disease and mild cognitive impairment during the COVID-19 lockdown,” *Frontiers in Psychiatry*, vol. 11, p. 590134, 2020.

[6] M. Maruta, H. Makizako, Y. Ikeda, H. Miyata, A. Nakamura, G. Han, and T. Tabira, “Association between apathy and satisfaction with meaningful activities in older adults with mild cognitive impairment: A population-based cross-sectional study,” *International Journal of Geriatric Psychiatry*, vol. 36, no.7, pp. 1065-1074, 2021.

[7] K. Ritchie, “Mild cognitive impairment: an epidemiological perspective,” *Dialogues in Clinical Neuroscience*, 2022.

[8] M. Kashefpoor, H. Rabbani, and M. Barekatain, “Supervised dictionary learning of EEG signals for

mild cognitive impairment diagnosis," *Biomedical Signal Processing and Control*, vol. 53, p. 101559, 2019.

[9] A. M. Alvi, S. Siuly, H. Wang, K. Wang, and F. Whittaker, "A deep learning based framework for diagnosis of mild cognitive impairment," *Knowledge-Based Systems*, vol. 248, p. 108815, 2022.

[10] F. Jamaloo, M. Mikaeili, and M. Noroozian, "Multi metric functional connectivity analysis based on continuous hidden Markov model with application in early diagnosis of Alzheimer's disease," *Biomedical Signal Processing and Control*, vol. 61, p. 102056, 2020.

[11] E. Andries, and R. Nikzad-Langerodi, "Dual-Constrained and Primal-Constrained principal component analysis," *Journal of Chemometrics*, e3403, 2022.

[12] J. Kevric, and A. Subasi, "The effect of multiscale PCA de-noising in epileptic seizure detection," *Journal of Medical Systems*, vol. 38, no. 10, pp. 1-13, 2014.

[13] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, "EEGdenoiseNet: A benchmark dataset for deep learning solutions of EEG denoising," *Journal of Neural Engineering*, vol. 18, no. 5, p. 056057, 2021.

[14] D. K. Barrow, and S. F. Crone, "Crogging (cross-validation aggregation) for forecasting—A novel algorithm of neural network ensembles on time series subsamples," *IEEE proceedings of 2013 International Joint Conference on Neural Networks (IJCNN)*, 2013, pp. 1-8.

[15] H. L. Vu, K. T. W. Ng, A. Richter, and C. An, "Analysis of input set characteristics and variances on k-fold cross validation for a Recurrent Neural Network model on waste disposal rate estimation," *Journal of Environmental Management*, vol. 311, p. 114869, 2022.

[16] A. Al-Qerem, F. Kharbat, S. Nashwan, S. Ashraf, and K. Blaou, "General model for best feature extraction of EEG using discrete wavelet transform wavelet family and differential evolution," *International Journal of Distributed Sensor Networks*, vol. 16, no. 3, p. 1550147720911009, 2020.

[17] S. H, Syed, and V. Muralidharan, "Feature extraction using Discrete Wavelet Transform for fault classification of planetary gearbox–A comparative study," *Applied Acoustics*, vol. 188, p. 108572, 2022.

[18] M. Ustundag, "A novel analog modulation classification: discrete wavelet transform-extreme learning machine (DWT-ELM)," *Bitlis Eren University Journal of Science*, vol. 10, no. 2, pp. 492-506, 2021.

[19] D. Benouioua, D. Candusso, F. Harel, and L. Oukhellou, "Multifractal analysis of stack voltage based on wavelet leaders: A new tool for PEMFC diagnosis," *Fuel Cells*, vol. 17, no. 2, pp. 217-224, 2017.

[20] E. Serrano, and A. Figliola, "Wavelet leaders: a new method to estimate the multifractal singularity spectra," *Physica A: Statistical Mechanics and its Applications*, vol.388, no.14, pp. 2793-2805, 2009.

[21] R. F. Leonarduzzi, G. Schlotthauer, and M. E. Torres, "Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia," *In 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 110-113.

[22] K. Gadhoumi, D. Do, F. Badilini, M. M. Pelter, and X. Hu, "Wavelet leader multifractal analysis of heart rate variability in atrial fibrillation," *Journal of Electrocardiology*, vol. 51, no. 6, pp. S83-S87, 2018.

[23] Z. Tan, and J. Chen, "Detecting stock market turning points using wavelet leaders method," *Physica A: Statistical Mechanics and its Applications*, vol. 565, p. 125560, 2021.

[24] Z. H. Zhou, "Ensemble learning," *In Machine learning*, Singapore: Springer, 2021, pp. 181-210. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-15-1967-3_8#citeas

[25] A. A. ABRO, "Vote-based: Ensemble approach," *Sakarya University Journal of Science*, vol. 25, no. 3, pp. 858-866, 2021.

[26] R. Salam, and A. R. M. T. Islam, "Potential of RT, Bagging and RS ensemble learning algorithms for reference evapotranspiration prediction using climatic data-limited humid region in Bangladesh," *Journal of Hydrology*, vol. 590, p. 125241, 2020.

[27] A. Saday, and I. A. Ozkan, "Classification of epileptic EEG signals using DWT-based feature extraction and machine learning methods," *International Journal of Applied Mathematics Electronics and Computers*, vol. 9, no. 4, pp. 122-129, 2021.

[28] P. Chen, and C. Pan, "Diabetes classification model based on boosting algorithms," *BMC Bioinformatics*, vol. 19, pp.1-9, 2018.

[29] S. Krishnaveni, and M. Hemalatha, "A perspective analysis of traffic accident using data mining techniques," *International Journal of Computer Applications*, vol. 23, no. 7, pp. 40-48, 2011.

[30] S. J. Ruiz-Gómez, C. Gómez, J. Poza, G. C. Gutiérrez-Tobal, M. A. Tola-Arribas, M. Cano, and R. Hornero, "Automated multiclass classification of spontaneous EEG activity in Alzheimer's disease and mild cognitive impairment," *Entropy*, vol. 20, no. 1, p. 35, 2018.

[31] J. Poza, C. Gomez, M. Garcia, M. A Tola-Arribas, A. Carreres, M. Cano, and R. Hornero, "Spatio-temporal fluctuations of neural dynamics in mild cognitive

impairment and Alzheimer's disease," *Current Alzheimer Research*, vol. 14, no. 9, pp. 924-936, 2017.

[32] S. Hadiyoso, and L. E. Tati, "Mild Cognitive Impairment Classification using Hjorth Descriptor Based on EEG Signal," *In 2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC)*, 2018, pp. 231-234.

[33] S. Hadiyoso, C. L. F. A. R. Cynthia, M. T. L. ER, and H. Zakaria, "Early detection of mild cognitive impairment using quantitative analysis of EEG signals," *IEEE proceedings of 2019 2nd International Conference on Bioinformatics, Biotechnology and Biomedical Engineering (BioMIC)-Bioinformatics and Biomedical Engineering*, 2019, pp. 1-5.

[34] M. Kashefpoor, H. Rabbani, and M. Barekatain, "Automatic diagnosis of mild cognitive impairment using electroencephalogram spectral features," *Journal of Medical Signals and Sensors*, vol. 6, no. 1, p. 25, 2016.