Int. j. inf. tecnol.
https://doi.org/10.1007/s41870-023-01676-y

ORIGINAL RESEARCH

# Utilizing graph Fourier transform for automatic Alzheimer's disease detection from EEG signals

**Ramnivas Sharma**[1] · **Hemant Kumar Meena**[1]

Received: 22 August 2023 / Accepted: 28 November 2023


**Abstract** Alzheimer's disease (AD) is rapidly increasing globally, poses a significant challenge in the field of neurodegeneration. A critical aspect is the development of biomarker tools to detect and monitor the disease's progression, especially in its early stages. Electroencephalogram (EEG) signal analysis holds promise for automated Alzheimer's diagnosis. However, existing methods frequently disregard the intricate functional interconnections within the brain during diverse activities, as they analyze each EEG channel in isolation. In response to these limitations, this research introduces a novel approach grounded in graph signal processing (GSP). In this paper, a new feature of graph Fourier transform (GFT) based on the concept of GSP has been introduced for the detection of Alzheimer. We have compared our work with the popular feature of discrete wavelet transform (DWT) and the existing features based methods. Our method provides 98.9% accuracy using decision tree (DT) classifiers on the Florida-based dataset, which is better than the existing state-of-art techniques for Alzheimer's detection.

**Keywords** EEG signals · Alzheimer's disease · Discrete wavelet transform (DWT) · Graph Fourier transform (GFT) · Machine learning classifiers

## 1 Introduction

Alzheimer's disease (AD) is a complex neurodegenerative condition that is the most common form of dementia. As it advances, it affects brain cells [1]. Alzheimer's disease is widely recognized neurodegenerative condition that leads to cognitive decline. It's the most studied nervous system disorder in the medical field. Despite extensive research, there's no cure to halt its progression. Hence, Early AD diagnosis is necessary for effective treatment and future patient care planning. However, various treatment options, including medication and non-medication approaches, can help manage AD symptoms at different stages, improving patients' quality of life. Alzheimer's disease is categorized into three stages: mild, moderate, and severe. As the disease progresses, the cognitive functions of patients gradually deteriorate [2]. Creating diverse biomarker tools is pivotal in detecting and monitoring the advancement of Alzheimer's, particularly during its initial phases. The electroencephalogram (EEG) signal analysis presents a promising avenue for automated identification of Alzheimer's disease. The preference for EEG signals in these systems stems from their cost-effectiveness and shorter time demands, setting them apart from alternative biomedical imaging modalities like CT (Computed Tomography), MRI (Magnetic Resonance Imaging), PET (Positron Emission Tomography), and fMRI (Functional Magnetic Resonance Imaging) [3]. The primary instigators of this neurological ailment are neuronal cell deterioration and dysfunction.

EEG is a diagnostic tool that reflects the electrical activity of cortical neurons in the brain, offering insights into brain-related physiological conditions [2]. Brain activity exhibits distinct rhythms categorized by frequency ranges: $\delta$ (0.5–4 Hz), $\theta$ (4–7 Hz), $\alpha$ (8–13 Hz), $\beta$ (13–30 Hz), and $\gamma$ ($\geq$ 30 Hz). These rhythms signify different mental states and functions. Under the hypothesis that resting EEG reflects overall brain

✉ Ramnivas Sharma
2020ree9534@mnit.ac.in

Hemant Kumar Meena
hmeena.ee@mnit.ac.in

[1] Department of Electrical Engineering, Malaviya National Institute of Technology, Jaipur, Rajasthan, India

Published online: 27 January 2024

state, specific frequency patterns may indicate abnormalities [3]. Thus, it's postulated that EEG time series from healthy individuals differ from those with neurodegenerative conditions like Alzheimer's or other disorders like epilepsy. Studies suggest that Alzheimer's has three notable effects on EEG: reduced complexity, signal slowing, and disruptions in synchrony. Research in spectral analysis shows increased delta and theta activity alongside decreased alpha and beta activity in AD patients, indicating a slowing of EEG signal frequency relative to age and alertness [3, 4].

Previous studies have shown that EEG abnormalities in AD reflect the anatomical and functional deficits of the cerebral cortex damaged by the disease. Kulkarni [4] combined spectral and complex features, achieving 94% accuracy with k-nearest neighbors (KNN). In a study by Fouad, I.A et al. [5], an Alzheimer's disease identification approach was presented. Results show that Naive Bayes and lagrangian support vector machine (LSVM) classifiers achieved the highest accuracy of 96.55% and 95.69%, and using a "ResNet-50" Convolutional Neural Network achieved 97.8261%. In [6], an long short term memory (LSTM)-based framework was proposed for early MCI detection. After assessing 20 LSTM models on a public MCI database, the best-performing model achieved 96.41% accuracy. In [7], a novel method is proposed using low-complexity orthogonal wavelet filter bank (LCOWFBs), HFD, and KFD achieved high classification accuracies of 98.5% (LCOWFBs-4) and 98.6% (LCOWFBs-6) with a cubic-support vector machine classifier and 10-fold cross-validation. Table 1 presents a compendium of research accomplishments within the domain of AD detection employing various methodologies.

From the above discussion traditional machine learning approaches encounter challenges when analyzing EEG signals, such as overlooking spatial relationships among electrodes, struggling with the brain's non-Euclidean geometry, and inadequately capturing temporal dynamics. GFT based on GSP is preferred for EEG signals due to its ability to naturally capture brain connectivity using graph structures, incorporate spatial arrangement and temporal dynamics, handle non-euclidean geometry, adapt to dynamic changes, and effectively denoise data. Therefore in this paper in addition to the machine learning model we incorporate the GFT for EEG signal analysis.

**Main contribution of work**

- Introduce a new graph signal-based feature such as Graph Fourier Transform (GFT) which capturing brain connectivity, spatial arrangement, temporal dynamics, non-Euclidean geometry.
- In this paper extracting attributes (mean, kurtosis, standard deviation, minimum, maximum, median) from discrete DWT and GFT and implemented on classifiers such as decision tree (DT), random forest (RF), logistic regression (LR), support vector machine (SVM), k-nearest neighbors (KNN) and extreme gradient boosting (XGBoost).
- Assessment of classification efficacy using six classifiers from distinct categories, fine-tuned with optimized hyper-parameters, for AD detection.

The subsequent sections of the paper follow this structure: In Sect. 2, we offer an overview of the dataset, the features employed in this study, and an outline of the essential stages of the proposed methodology. In Sect. 3, we present the results of a performance comparison between our method and other approaches. Finally, the paper is concluded in Sect. 4.

## 2 Proposed methodology

The Alzheimer's Disease diagnostic system comprises three sequential phases: firstly, data collection and preprocessing, followed by the extraction of features from the EEG data, and ultimately, the classification process, as depicted in Fig. 1. Algorithm 1 provides a detailed algorithmic representation of the methodology. In the following sections, we offer detailed insights into the dataset, feature extraction process and the classifiers used in this system.

**Table 1** Summary of work for Alzheimer's disease classification using different techniques

| Refs. | Method | Observation | Accuracy |
|---|---|---|---|
| [1] | SVM, KNN | Information on the sensitivity and specificity of the proposed method is missing | 94% |
| [3] | GRU, LSTM SVM, KNN | Performance of GRU is compared with LSTM, SVM and KNN | 96.91% |
| [4] | KNN | Only 1 ML model is used No comparison performed | 96% |
| [6] | LSTM | Does not compare the proposed LSTM-based framework with other existing models | 96.41% |
| [8] | ML classifiers | Specific limitations of the proposed LRFB model | Outperformed models by 55.18% |
| [9] | BayesNet, Kernel Support vector machine | Does not compare the proposed approach with other existing methods | 77.76% |

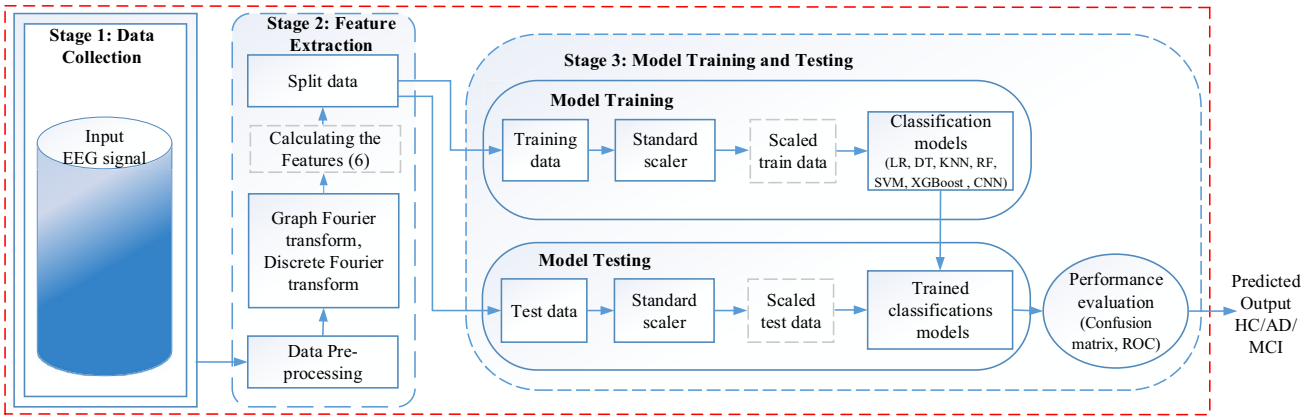


**Fig. 1** Schematic diagram depicting the structure of the proposed approach

## 2.1 Dataset

This study used a dataset [5, 10] gathered by researchers from Florida State University, utilizing a 19-electrode recording setup with the Biologic Systems Brain Atlas III Plus workstation, following the international 10–20 system. This dataset encompassed recordings from four distinct groups labeled A, B, C, and D. Groups A and C underwent recordings with their eyes open while maintaining visual fixation, whereas groups B and D were recorded with their eyes closed. The participant composition encompassed twenty-four healthy control (HC) elderly individuals for groups A and B. Conversely, groups C and D consisted of 24 subjects diagnosed as probable Alzheimer's disease (AD) patients. In our research, we conducted separate experiments on two groups: one with open eyes and the other with closed eyes. This EEG dataset is openly accessible in [10].

## 2.2 Graph Fourier transform (GFT)

The GFT [11, 12] finds relevance in EEG signal analysis, extending the classical Fourier Transform to irregular structures like brain connectivity networks.EEG signals mirror brain activity and can be expressed as graph signals, using electrodes as nodes and connections as edges. GFT dissects these signals into graph eigenfunctions, uncovering inherent frequency patterns in brain dynamics. This decomposition is based on the graph's Laplacian matrix. These eigenfunctions form a basis for analysis, revealing patterns and characteristics of the graph data. The fundamental process of acquiring the GFT is depicted in Fig. 2.

The graph Laplacian matrix $\mathcal{L}$ [11] is defined as

$$\mathcal{L} = D - W \tag{1}$$

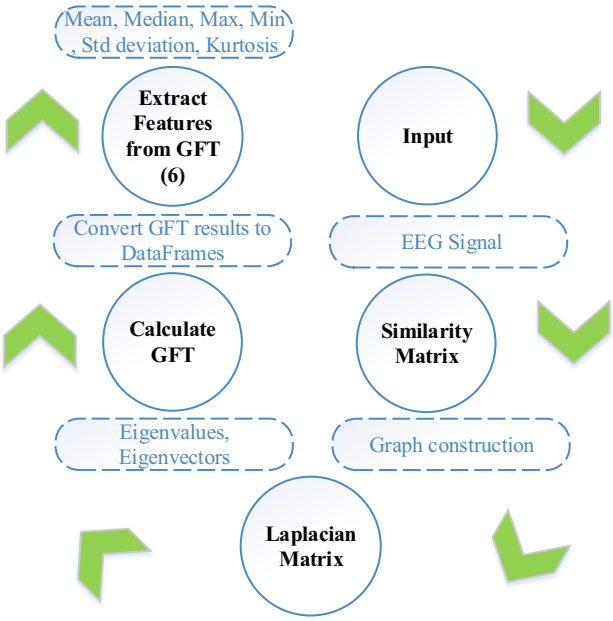


**Fig. 2** Procedure for obtaining the graph Fourier transform of a signal

where $D \rightarrow$ diagonal matrix, $W \rightarrow$ adjacency matrix (similarity matrix). The normalized graph Laplacian [11] is represented as

$$\tilde{\mathcal{L}} = I_N - \boldsymbol{D}^{-1/2}\boldsymbol{W}\boldsymbol{D}^{-1/2} \tag{2}$$

where $I_N$ an identity matrix of size $N$ and $\boldsymbol{D}$ denotes diagonal matrix. GFT of graph signal $\mathbf{x}$ on graph $G$ is expressed as $\tilde{\mathbf{x}} = \boldsymbol{U}^{\boldsymbol{T}}\mathbf{x}$ while the inverse graph Fourier is given by $\mathbf{x} = \boldsymbol{U}\tilde{x}$ [11, 12]. where $\boldsymbol{U}$ comprise of eigenvectors $\{u_l\}, l = 1, \ldots, N$.

**Algorithm 1** Data processing and classification pipeline

```
Require: Directories A, B, C, D
Import necessary libraries
Function Input Processing dir
for each file in os.ls(dir) do
    Construct path: p ← join(dir, file)
    Read CSV into df
end for
Function ObtainGFT(dataset):
Calculate similarity: S ← sim(dataset)
Create graph: G ← graph(S)
Calculate Laplacian: L ← lap(G)
Eigenvalues, eigenvectors: λ, u ← eig(L)
End Function
Function ExtractFeatures(df_g): features
for each row in df_g do
    Calculate various statistics for the row (e.g., mean, std, median, min, max, kurtosis)
end for
    End Function
    Split data, Scale features and Define classifiers
    for each clf in classifiers do
        Initialize a GridSearchCV classifier with the model and hyperparameters
        Fit grid_clf on X_train, y_train
        best_clf ← best(grid_clf)
        y_pred ← predict(best_clf, X_test)
        Calculate accuracy, precision, recall, and F1-score
        Display clf, best_params, metrics
    end for
```

### 2.3 Discrete wavelet transform (DWT)

The DWT [5, 13] is a signal processing technique that decomposes a signal into different frequency bands. In this paper, we used $4^{th}$ level wavelet transform function 'db4' developed by Daubechies on EEG signals, which involves a series of low-pass ($A_{j,k}$, approximation coefficients) and high-pass ( $D_{j,k}$, detail coefficients) filtering operations that successively decompose the signal into different scales or levels. In this article, we acquire various statistical characteristics derived from the DWT coefficients and the GFT. These features encompass parameters like the mean, maximum, minimum, kurtosis, median, and standard deviation. The function $\phi_{j,k}[n]$ is determined by the low-pass filter, while the wavelet function $\psi_{j,k}[n]$ is determined by the high-pass filter, as visually depicted. The DWT coefficients are defined by the given below equation:

$$\psi_{j,k}[n] = 2^{j/2} H(2^j n - k) \tag{3}$$

$$\phi_{j,k}[n] = 2^{j/2} L(2^j n - k) \tag{4}$$

$$D_{j,k} = \sum_{n=0}^{N-1} x[n] \cdot \psi_{j,k}[n] \tag{5}$$

$$A_{j,k} = \sum_{n=0}^{N-1} x[n] \cdot \phi_{j,k}[n] \tag{6}$$

where $D_{j,k}$ and $A_{j,k}$ are the wavelet and scaling coefficients respectively.

## 3 Results and discussion

In our investigation, we partitioned the dataset into two sets: SET A & C comprised the eyes-open group, while SET B & D constituted the eyes-closed group. After applying rigorous pre-processing techniques to eliminate noise and artifacts, the EEG signals were further analyzed using three methods: 1. Utilizing the DWT, 2. Employing the GFT, and 3. Calculating statistical features from both GFT and DWT. The results obtained from these approaches were subsequently juxtaposed against various machine learning algorithms. For each group, whether eyes open or closed, a collection of traditional features encompassing parameters such as mean, maximum, minimum, median, standard deviation, and kurtosis were extracted following the signal decomposition. The final step involved classifying two categories: healthy subjects and those with Alzheimer's. This categorization was achieved using an array of machine learning classifiers, including support vector machine (SVM), k-nearest neighbors (KNN), decision tree (DT), random forest (RF), logistic regression (LR), and Extreme gradient boosting

**Table 2** Parameters of different machine learning classifiers

| Models | Parameters |
|---|---|
| Decision Tree (DT) | criterion: entropy, max_depth: 5 |
| Random Forest (RF) | max_depth: None, n_estimators: 50 |
| Logistic Regression (LR) | C: 10, penalty: l2 |
| Support Vector Machine (SVM) | C: 10, kernel: rbf |
| k-Nearest Neighbors (KNN) | n_neighbors: 5 |
| Extreme Gradient Boosting (XGBoost) | max_depth: 5, n_estimators: 200 |

**Table 3** Accuracy of various classifiers on the SET A & C subset (open eyes group) of EEG data

| Classifiers | SET A & C using DWT and GFT | | | | | XGBoost |
|---|---|---|---|---|---|---|
| | DT | RF | LR | SVM | KNN | |
| Acc.(%) DWT | 89.36 | 90.43 | 87.94 | 89.54 | 87.59 | 89.54 |
| Acc.(%) GFT | 89.89 | 91.49 | 86.52 | 92.20 | 90.07 | 91.84 |

**Table 4** Accuracy of different classifiers in the context of SET B & D (closed eyes group) of data

| Classifiers | SET B & D using DWT and GFT | | | | | XGBoost |
|---|---|---|---|---|---|---|
| | DT | RF | LR | SVM | KNN | |
| Acc.(%) DWT | 98.29 | 98.54 | 98 | 98.29 | 97.07 | 98.29 |
| Acc.(%) GFT | 98.9 | 98.29 | 98.04 | 98.04 | 97.32 | 98.29 |

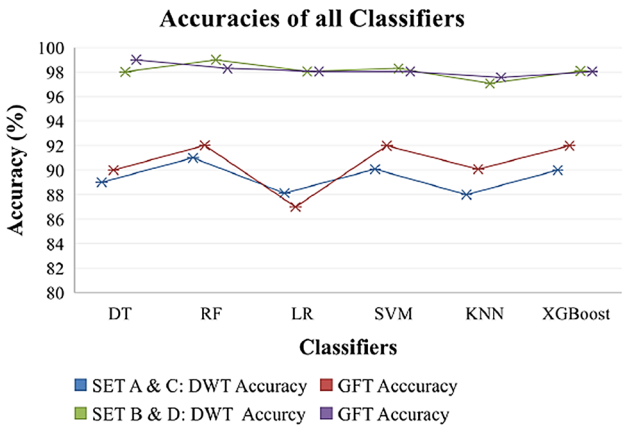


**Fig. 3** Accuracies of all proposed classifiers on EEG dataset

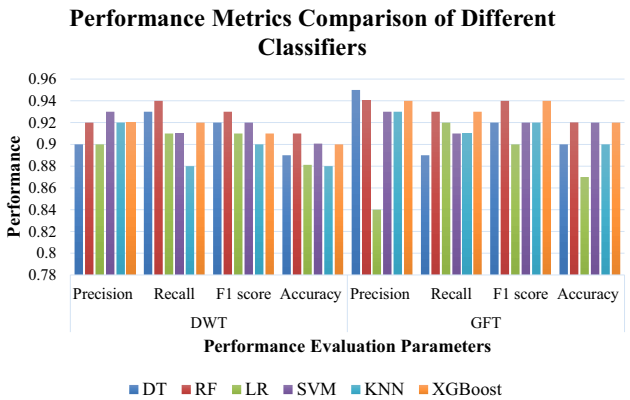


**Fig. 4** Comparison of performance metrics on SET A & C Curve (open eyes group)

(XGBoost). Table 2 displays the parameters for all machine learning classifiers. Tables 3 and 4 present the outcomes derived from diverse classifiers, utilizing both GFT and DWT techniques, for SET A & C (open eyes group) and

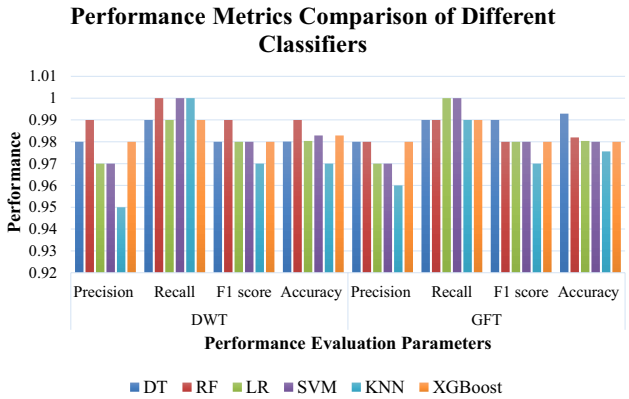


**Fig. 5** Comparison of performance metrics on SET B & D (closed eyes group)

SET B & D (closed eyes group), respectively. A comparison between Tables 3 and 4 results underscores the enhancement in classifier performance achieved by incorporating GFT parameters as informative features across all decomposition modes. Notably, the DT classifier achieved the most promising outcome with an accuracy of 98.9%. The accuracy outcomes of different classifiers and methods for DWT and GFT-based feature sets are visually depicted in Fig. 3. Additionally, an evaluation of performance metrics for various classifiers, accompanied by their accuracy rates for both eyes open and eyes closed groups, is illustrated in Figs. 4 and 5 respectively. Table 5 summarizes previous investigations that employed EEG signals to diagnose AD. In contrast to earlier studies using many statistical attributes, our research achieves positive classification results with just six statistical characteristics and various classifiers.

## 4 Conclusion and future scope

This study investigates the potential of EEG as a novel method for early-stage Alzheimer's disease detection. The

**Table 5** Comparison of the proposed method with existing methods

| Refs. | EEG dataset | Class number | Classifiers | Features | Accuracy% |
|---|---|---|---|---|---|
| [1] | 50 HC + 50 AD | 2 | SVM | Spectral properties | 94 |
| [4] | 50 HC + 50 AD | 2 | KNN | Complexity properties | 96 |
| [5] | 24 HC+24 AD | 2 | Naive Bayes, LSVM, "ResNet-50" CNN | DWT+ statistical features | 96.55 , 95.69 , 97.82 |
| | 7 MCI, 59 AD, 102 NC | 3 | Naive Bayes, LSVM | | 96.55 , 94.52 |
| [6] | 16 HC, 11 MCI | 2 | LSTM | – | 96.41 |
| [7] | AD-12 and NC-11 | 2 | SVM, LCOWFBs-6 | HFD, KFD | 98.6 |
| [14] | 11 HC, 11 AD | 2 | FuzzyEn analysis | Lilliefors test, sample entropy | 96.36 |
| [15] | 35 HC, 31 Mild AD, 20 moderate AD | 3 | SVM, KNN | DWT, EMD, Hjorth parameters | 97.4 |
| [16] | 37 MCI, 52 HC, 52 AD | 3 | CNN | Extract the RGB image by CWT | 98.90 |
| [17] | 11 HC, 11 MCI | 2 | SVM, KNN | SWT | 96.94 |
| [18] | 64 HC, 64 MCI, 64 AD | 3 | CNN | D-PSD | 82.30 |
| Proposed method | 24 HC, 24 AD | 2 | SVM | GWT, DWT and statistical parameters | 98.29 |
| | | | DT | | **98.9** |
| | | | LR | | 98.04 |
| | | | RF | | 98.54 |
| | | | KNN | | 97.32 |
| | | | XGBoost | | 98.29 |

Bold value reflects the highest accuracy among all the classifiers

findings suggest that a combined approach involving both GFT and wavelet features leads to improved classification rates and diagnostic accuracy. The results indicate this combined approach outperforms individual methods in terms of accuracy, reaching up to 98.9%. Future research can focus on improving the accuracy and precision of Alzheimer's disease detection using EEG signals by refining the application of the GFT and exploring more advanced signal processing techniques. This may involve optimizing feature extraction methods or incorporating deep learning approaches. The future scope encompasses refining accuracy, integrating multi-modal data, conducting longitudinal studies, developing real-time monitoring systems, extensive clinical validation, personalized medicine approaches, ethical considerations, and global implementation, collectively aiming to advance early Alzheimer's disease detection and personalized patient care using EEG-based GFT analysis.

**Author contributions** RS: original draft, software, review and editing of the paper, and results obtained.

**Funding** No Funding is involved.

**Data availability** Publicly available data has been referred to in this paper.

**Declarations**

**Conflict of interest** No competing financial or interpersonal conflicts.

**Ethical approval** Not involve any studies with animals or humans.

**Consent to participate** Human participants are not involved.

## References

1. Bairagi V (2018) EEG signal analysis for early diagnosis of alzheimer disease using spectral and wavelet based features. Int J Inf Technol 10(3):403–412
2. Mavrodaris A, Powell J, Thorogood M (2013) Prevalences of dementia and cognitive impairment among older people in subsaharan Africa: a systematic review. Bull World Health Organ 91:773–783
3. Alvi AM, Siuly S, Wang H, Wang K, Whittaker F (2022) A deep learning based framework for diagnosis of mild cognitive impairment. Knowl-Based Syst 248:108815
4. Kulkarni N (2018) Use of complexity based features in diagnosis of mild Alzheimer disease using EEG signals. Int J Inf Technol 10(1):59–64
5. Fouad IA, Labib FE-ZM (2023) Identification of Alzheimer's disease from central lobe EEG signals utilizing machine learning and residual neural network. Biomed Signal Process Control 86:105266
6. Alvi AM, Siuly S, Wang H (2022) A long short-term memory based framework for early detection of mild cognitive impairment

from EEG signals. IEEE Trans Emerg Topics Comput Intell 7(2):375–388
7. Puri DV, Nalbalwar SL, Nandgaonkar AB, Gawande JP, Wagh A (2023) Automatic detection of Alzheimer's disease from EEG signals using low-complexity orthogonal wavelet filter banks. Biomed Signal Process Control 81:104439
8. Shastry KA, Sattar SA (2023) Logistic random forest boosting technique for Alzheimer's diagnosis. Int J Inf Technol 15(3):1719–1731
9. Shaikh TA, Ali R (2022) Enhanced computerised diagnosis of Alzheimer's disease from brain mri images using a classifier merger strategy. Int J Inf Technol 14(4):1791–1803
10. Pineda AM, Ramos FM, Betting LE, Campanharo AS (2020) Quantile graphs for EEG-based diagnosis of Alzheimer's disease. PLoS One 15(6):e0231169
11. Shuman DI, Narang SK, Frossard P, Ortega A, Vandergheynst P (2013) The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process Mag 30(3):83–98
12. Stanković L, Daković M, Sejdić E (2019) Introduction to Graph Signal Processing. In: Stanković L, Sejdić E (eds) Vertex-Frequency Analysis of Graph Signals. Signals and Communication Technology. Springer, Cham. https://doi.org/10.1007/978-3-030-03574-7_1
13. Amin HU, Malik AS, Ahmad RF, Badruddin N, Kamel N, Hussain M, Chooi W-T (2015) Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australas Phys Eng Sci med 38:139–149
14. Simons S, Espino P, Abásolo D (2018) Fuzzy entropy analysis of the electroencephalogram in patients with Alzheimer's disease: is the method superior to sample entropy? Entropy 20(1):21
15. Safi MS, Safi SMM (2021) Early detection of Alzheimer's disease from EEG signals using hjorth parameters. Biomed Signal Process Control 65:102338
16. Huggins CJ, Escudero J, Parra MA, Scally B, Anghinah R, De Araújo AVL, Basile LF, Abasolo D (2021) Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing. J Neural Eng 18(4):046087
17. Yin J, Cao J, Siuly S, Wang H (2019) An integrated mci detection framework based on spectral-temporal analysis. Int J Autom Comput 16:786–799
18. Amini M, Pedram MM, Moradi A, Ouchani M (2021) Diagnosis of Alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using EEG signal. Comput Math Methods Med 2021:5511922