



Graphical representation learning-based approach for automatic classification of electroencephalogram signals in depression

Surbhi Soni^a, Ayan Seal^{a,*}, Anis Yazidi^{b,d,c}, Ondrej Krejcar^{e,f}

^a PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, 482005, India

^b Department of Computer Science, OsloMet – Oslo Metropolitan University, Oslo, Norway

^c Department of Plastic and Reconstructive Surgery, Oslo University Hospital, Oslo, Norway

^d Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway

^e Center for Basic and Applied Research, Faculty of Informatics and Management, University of Hradec Kralove, Hradecka 1249, Hradec Kralove, 50003, Czech Republic

^f Malaysia Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100, Kuala Lumpur, Malaysia



ARTICLE INFO

Keywords:

Node2vec
Electroencephalography
Graph-level fusion
Feature-level fusion
Decision-level fusion

ABSTRACT

Depression is a major depressive disorder characterized by persistent sadness and a sense of worthlessness, as well as a loss of interest in pleasurable activities, which leads to a variety of physical and emotional problems. It is a worldwide illness that affects millions of people and should be detected at an early stage to prevent negative effects on an individual's life. Electroencephalogram (EEG) is a non-invasive technique for detecting depression that analyses brain signals to determine the current mental state of depressed subjects. In this study, we propose a method for automatic feature extraction to detect depression by first constructing a graph from the dataset where the nodes represent the subjects in the dataset and where the edge weights obtained using the Euclidean distance reflect the relationship between them. The Node2vec algorithmic framework is then used to compute feature representations for nodes in a graph in the form of node embeddings ensuring that similar nodes in the graph remain near in the embedding. These node embeddings act as useful features which can be directly used by classification algorithms to determine whether a subject is depressed thus reducing the effort required for manual handcrafted feature extraction. To combine the features collected from the multiple channels of the EEG data, the method proposes three types of fusion methods: graph-level fusion, feature-level fusion, and decision-level fusion. The proposed method is tested on three publicly available datasets with 3, 20, and 128 channels, respectively, and compared to five state-of-the-art methods. The results show that the proposed method detects depression effectively with a peak accuracy of 0.933 in decision-level fusion, which is the highest among the state-of-the-art methods.

1. Introduction

Depressive disorders are characterized by sadness, loss of interest or pleasure, disturbed sleep or appetite, a sense of guilt and hopelessness, poor concentration, and can even lead to suicide in its most severe form. According to the World Health Organization, 322 million people worldwide suffer from depression [1]. Thus, diagnosing at an early stage is critical to protect patients from the severe and irreversible consequences of depression. It has three levels, namely mild, moderate, and acute. Doctors diagnose the same based on classification criteria such as the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition, Beck Depression Inventory, which are a questionnaire containing a set of questions score of which determines the level of depression. Also, the

interactive sessions between the patients and health practitioners play an imperative role in detecting depression. However, these sessions may not always produce the desired results because it is dependent on the practitioner's knowledge and skill in dealing with depressive patients. Moreover, patients suffering from a mental disorder are hesitant to admit that they require treatment because they are afraid of being labeled mentally ill. As a result, clinical treatment is ineffective in assisting depressive patients in receiving proper treatment, resulting in further deterioration of the illness and its detrimental effect on an individual's life. Alternative approaches to effectively detect and diagnose depression are required to overcome the drawbacks of clinical methods. The advancement of sensor technology and communication systems has resulted in the development of electronic devices that are critical in

* Corresponding author.

E-mail address: ayan@iitdmj.ac.in (A. Seal).

monitoring an individual's health conditions. These devices are commonly employed to investigate the activity of the human brain in techniques such as electroencephalography (EEG), magnetoencephalography, magnetic resonance imaging, and functional magnetic resonance imaging. Out of these, EEG is a portable and non-invasive technique that uses electrodes attached to the scalp to evaluate the electrical activity of the brain. This electrical activity is represented on EEG recordings by a wavy line, which is then used by physicians and scientists to study brain functions and diagnose neurological disorders. The electrodes employed are generally useable and are labeled as F, P, O, and T, which represent the areas of the brain from which the signals are acquired, such as the frontal, parietal, occipital, and temporal lobes. The electrode in the midline of the brain is designated as z.

EEG evaluates the functioning of the brain and is used for diagnosis of illness like epileptic seizure [2], Parkinson's Disease [3], Schizophrenia [4], Stress Detection [5] and Sleep Disorder [6]. The use of EEG signals in the detection of depression is a new area of study. It begins with the extraction of EEG signals using an electroencephalogram device, followed by pre-processing, which removes disturbances in EEG signals caused by environmental changes, commonly referred to as artifacts. Following the removal of artifacts, the important features of EEG signals are extracted and fed into the classification algorithms, that determine whether or not the subject is depressed. Because of nonlinear properties EEG data, extracting useful features from it is difficult. The features and behaviour of EEG signals cannot be clearly explained using time, frequency, or time-frequency analysis. As a result, it's critical to locate stable features and create accurate models with high classification performance [7]. Hence the feature extraction procedure entails the manual extraction of handcrafted features, which is tedious and labor-intensive. The proposed method is based on an automated feature extraction process based on graphical representations, which overcomes the problem of manual feature extraction.

1.1. Motivation and contributions

EEG processing entails signal processing and analysis, which begins with the extraction of useful features in the time, frequency, and time-frequency domains. This process takes a long time and a lot of human effort; additionally, there are no fixed global biomarker features for detecting depression. As a result, the work aims to reduce the manual effort required for feature extraction and propose an automatic method for detecting depressive patients using graphical representation learning. The contributions of this work are as follows:

- The proposed method aims to develop a predictive model that automatically extracts features from EEG signals using graph representation learning followed by the classification of the depressed and healthy subjects using machine learning algorithms.
- The proposed method develops three novel approaches graph-level, feature-level, and decision-level data fusion techniques that provides a way to combine the features extracted from each channel of the dataset that are obtained after applying the Node2vec algorithm in order to obtain a greater predictive power. As far as these fusion techniques are concerned, they are mainly exploiting the Node2vec algorithmic framework to carry out the data fusion procedure.
- The Node2vec algorithmic framework is adopted that computes a vector representation of a node based on random walks in the graph. These vectors corresponding to each node acts as features and is fed into machine learning algorithms to classify depressed and healthy subjects.
- The proposed method is executed on three datasets consisting of 3 channels, 20 channels, and 128 channels respectively. It is then compared with five state-of-the-art methods using different evaluation metrics, namely accuracy, sensitivity/recall, specificity, precision, and ROC curve. Empirical results illustrate that the decision-level-based proposed method outranks some existing approaches.

The paper's structure is as follows, Section II discusses a literature review of existing depression-related works. The proposed work is described in detail in Section III, which includes a detailed explanation of the Node2vec method. The results and discussions are presented in Section IV, followed by the conclusion and future scope in Section V.

2. Related work

A substantial amount of research had been conducted in the field of depression detection using EEG signals. Hinrikus et al. [8] presented a new method for analyzing EEG signals based on the frequency spectrum, assuming that beta-band played an important role in the detection of depression. The Spectral Asymmetry Index (SASI) was exploited as a promising measure to detect depressed patients and was found positive in the depressed groups and negative in the healthy group. However, no statistically significant features related to hemispherical asymmetry had been observed. Grin-Yatsenko et al. [9] employed the Independent Component Analysis method to decompose raw EEG data extracted from multiple channels into separate components for each eye open and eye closed state. Every component's and each subject's frequency band power spectra were computed. A comparison of the power spectrum in theta, alpha, and beta bands reported exceptionally high alpha power during resting state in depressed patients, which was contradicted by Stewart et al. [10], whose findings showed low alpha activity in all parts of the brain during the depression. Gerard E. Bruder et al. [11] used dichotic test demonstrations in which external stimuli such as Fusion words and complex tones were delivered and hemisphere Asymmetry [12] was computed to classify depressed and healthy subjects. It claimed that healthy men had greater hemispheric asymmetry than women in perceiving tones and words, whereas depressed patients showed no gender difference. Other measures, however, were required to further depict the gender differences in depression. Discrete Wavelet transform was employed by Puthankattil and Joseph [13] to decompose the EEG signal data extracted from 15 depressed and 15 healthy subjects into different frequency bands and Parseval's method [14], was applied to calculate energy present at various decomposition levels. The two-layered feed-forward neural network with twenty neurons was adopted to classify the depressed and healthy subjects based on relative wave entropy and signal entropy features. However, it was a preliminary study that accepted a small number of features and a conventional algorithm on a smaller dataset. Ahmadlou et al. [15] found the Higuchi Fractal Dimension of the left, right, and overall frontal brain beta sub-bands effective in distinguishing between depressed and healthy subjects when passed to the enhanced probabilistic neural network for classification. The research findings were preliminary because it experimented on a smaller dataset and additional data was required to draw definite conclusions. Bachmann et al. [16] in their findings used SASI and HFD to differentiate between depressive and healthy subjects and suggested that the HFD alone did not perform well in the frontal region. According to Khoa et al. [17], HFD served as a successful discriminator in detecting depressed subjects in the parietooccipital region and not in the frontal region as it is more sensitive to noise because of its high variability. Hosseiniard et al. [18] findings concluded that non-linear features such as Correlation Dimension, Higuchi Fractal Dimension, and Lyapunov exponent were more effective in detecting depression; however, combining both linear and non-linear features can improve accuracy even further. Faust [19] findings suggested that the Probabilistic Neural Network was capable of distinguishing between depressed and healthy subjects based on non-linear features such as Sample Entropy, Approximate Entropy, Renyi's Entropy, and Bispectral Entropy. However, the method used redundant features as the feature selection technique was not involved and high accuracy could have resulted from overfitting. Acharya et al. [20] extracted fifteen non-linear features and then ranked them using the t-test, and the features with the highest rank in both the left and right hemispheres were then combined to form the Depression Diagnostic

Index to effectively discriminate between the depressive and healthy groups. However, because there is no proof of the relationship of features used as a combination in the index's formulation, its use is questionable. Cai et al. [21] suggested a method in which EEG data was collected from three electrodes, namely Fp1, Fz, and Fp2 followed by its division into alpha, beta, gamma, and theta frequency bands. Twenty-seven linear and non-linear features were extracted from these frequency bands and fed into four classifiers, namely K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Artificial Neural Network and, Deep Belief Neural Network (DBNN). The results indicated that the absolute power of the beta sub-band combined with the DBNN Classifier provided an accuracy of 78.24% in distinguishing between the depressed and healthy groups. However, there was a need to increase data collection to gradually improve accuracy. Mumtaz et al. [22] extracted features from the EEG dataset consisting of nineteen channels which included the powers of different frequency bands and alpha interhemispheric asymmetry. Receiver Operating Criterion (ROC) [23] was used to select the most significant features to avoid irrelevant and redundant features and the passed for classification. The findings suggested that alpha interhemispheric asymmetry was the most promising characteristic when combined with an SVM Classifier. Liao et al. [24] developed a spectral-spatial feature extractor-based method that first filters raw multi-channel EEG signals into frequency bands and then transformed the original signals from each sub-band to an optimal feature vector space by reducing it into a lower-dimensional feature vector space. The resulting vector space provided the highest classification accuracy with the SVM classifier to distinguish between the depressed and healthy subjects. Sharma et al. [25] decomposed raw EEG signal from three channels into seven wavelet sub-bands using three Channel Orthogonal Wavelet Filter Bank. The logarithmic norm feature was extracted from each wavelet sub-band and fed into the support vector machine to differentiate between the depressed and healthy subjects. Cai et al. [26] conducted a study on 213 subjects, out of which 92 were depressed and 121 were normal and collected their EEG signals through a pervasive three-electrode EEG collector. The signals were denoised using a Finite Impulse Response Filter, then features were extracted and classified into three categories: time-domain features, frequency domain features, and non-linear features. To select the most important features the Minimal-Redundancy-Maximum-Relevancy method was adopted, and the results indicated that Absolute power theta in combination with the KNN classifier produced the best the highest accuracy of 76.98% in differentiating between the two groups, which was higher than previous studies. Byun et al. [27] used a machine learning approach to classify healthy and depressed subjects using Heat Rate Variability demonstration to evaluate cardiac autonomic regulation connection with depression. by extracting twenty HRV features—13 linear, five nonlinear, and two Poincaré plot from electrocardiogram (ECG) recordings of 37 depressed and 41 healthy subjects. The results suggested that SVM classifier distinguished the two groups with 74% accuracy. Cai et al. [28] presented a multimodal approach by fusing features extracted from three-channel EEG signals recorded in response to different modalities of positive, negative, and neutral audio stimulation. The extracted features combined with the KNN classifier yielded an accuracy of 86.98%. The studies mentioned above demonstrate the use of handcrafted features to detect depression that requires a lot of human effort. Our proposed method focuses on automated feature extraction using graphical representation hence reducing human effort in the feature extraction phase.

3. Methods and materials

The primary goal of this research is to propose a computer-aided automatic method for detecting depression using a graphical approach to feature extraction, thereby avoiding the use of handcrafted features that require human effort.

3.1. Feature extraction using Node2vec

A feature is a collection of variables that, when combined, best represent a larger set of data. Feature extraction is directly related to dimensionality reduction because it reduces the data set passed for training while accurately representing the original data set without sacrificing critical information. Node2vec [29] is an algorithmic framework that learns feature representation in the form of network nodes and is used for feature extraction from EEG data. It learns a node's dense representation in such a way that when it is plotted in a low-dimensional space, the nodes near it in that space are also its neighbors in the actual formation structure.

The Node2vec algorithm is divided into three steps. The first step is to generate a graph from the dataset, with nodes representing each subjects in the dataset and edges representing their Euclidean distance from one another. In the second step, for each node in a graph, second-order random walks are generated to form a sentence, which is a collection of node ids. A corpus, which is a collection of all sentences, is created. The third step is to generate node embedding, which is accomplished by passing the corpus to the skip-gram model, which treats each node id as a unique word/token in a dictionary to calculate the embedding vectors for each node, which are represented in a two-dimensional plane using the T-SNE algorithm, as shown in Fig. 1. These node embeddings serve as important features that will be used for classification in the future. Random walks can be thought of as a walker traversing the edges of a graph, deciding where to go next, and then moving on to the next step. The majority of existing random walk measures are based on the first-order Markov model, in which the next random walk step is determined solely by the current node. However, it is ineffective in many real-world applications, and experimental results show that second-order measures outperform their first-order counterparts in a variety of applications [30]. As a result, second-order random measures are used, which also consider the previously visited node and efficiently explore the various neighborhoods of a given node using hyper-parameters, which are as follows:

- **Return parameter (m):** The parameter controls the likelihood of the random walk returning to a previously visited node. The greater the value of m , the less likely it is that the random walk will return to a node. An m value less than one indicates that the random walk tends to return to a previously visited node in order to keep it close to the start node.
- **Inout parameter (o):** The parameter instructs the random walk to explore inward and outward from a specified source node. If o is less than one, a random walk will tend to stay close to the source node, demonstrating breadth first search behavior. As a result, the random walk includes nodes that are close to the source node, encouraging inward exploration. If o is greater than one, a random walk will tend to move away from the source node, demonstrating depth first search behavior to promote outward exploration.
- **Embedding dimension (d):** it denotes size of the embedding dimension.
- **Length of a random walk (l):** it denotes the number of nodes in a walk.
- **Walks per node (r):** it denotes the number of random walks for each node

The steps involved in the Node2vec algorithm are enlisted below.

1. Construct an undirected and weighted complete graph $G(V, E)$ from the EEG dataset, where V represents the set of nodes that depicts the subjects in the dataset and E is the set of weighted edges that reflect the distance between two nodes computed using euclidian distance.
2. Save the weights of all nodes to their neighbors, and calculate the transition probabilities between each node and its neighbors with the help of hyperparameters m and n using Eqs. (1) and (2), respectively.

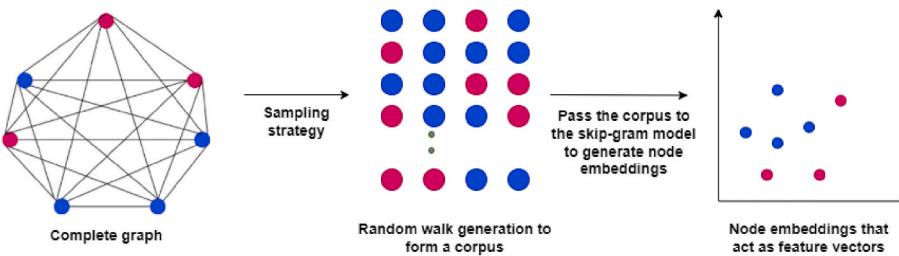


Fig. 1. Node2vec method of feature representation.

$$\pi_{tx} = \alpha_{mo}(u, x) \cdot w_{tx}$$

$$\alpha_{mo}(u, x) = \begin{cases} 1/m & \text{if } d_{ux} = 0 \\ 1 & \text{if } d_{ux} = 1 \\ 1/o & \text{if } d_{ux} = 2 \end{cases} \quad (2)$$

where t and x are nodes V_t and V_x respectively and $w_{t,x}$ reflects the weight between the two nodes. Consider Fig. 2 where a random walk that has just visited edge (u, t) and now resides at node t . The walk must now decide what to do next, so it computes and analyzes the transition probabilities π_{tx} on edges (t, x) leading from t . Here d_{ux} denotes the shortest path distance between nodes u and x .

3. Obtain random walks for a node u of fixed length 1 by generating nodes c_i using Eq. (3).

$$P(c_i = x | c_{i-1} = t) = \begin{cases} \frac{\pi_{tx}}{Z} & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where c_i denotes the i th node of the random walk starting from source node u and $c_{(i-1)}$ is the previously visited node, π_{tx} is the unnormalized transition probability between nodes t and x , and Z denotes the normalizing factor which is a sum of all transition probabilities.

4. Train all the walk paths of each node using a skip-gram model and then obtain the node embeddings of dimension d which act as relevant features.

3.2. Classification

In this study, a 10-fold cross-validation technique is adopted to

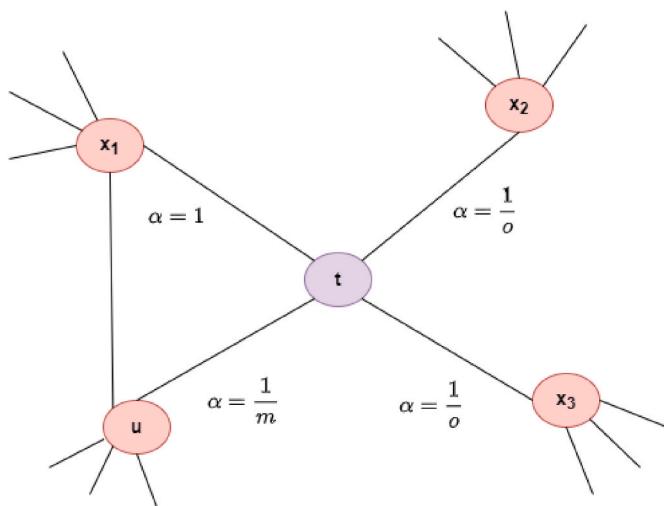


Fig. 2. Illustration of the random walk procedure in Node2vec method. The walk just transitioned from u to t and is now evaluating its next step out of node t to the neighboring nodes x_1, x_2 and x_3 . Search biases are shown by edge labels.

evaluate the performance of the proposed method, where the original dataset is randomly divided into ten equal-size subsamples [31]. A single subsample from the ten is kept as validation data for testing the method, while the remaining nine are used as training data. The cross-validation process is then performed ten times (the folds), with each of the ten subsamples serving as validation data exactly once. The evaluation metrics obtained from the 10 folds is then averaged (or otherwise combined) to estimate the performance of the proposed method. Here, six classifiers, namely Decision Tree (DT), KNN, XGBoost, SVM, Logistic Regression (LR), and Linear Discriminant Analysis (LDA), are adopted. These classifiers take feature vectors obtained from the validation set using the Node2vec algorithm separately and predict categories, namely depressed and healthy subjects. The classification algorithms along with their various parameters are listed in Table 1.

3.3. Evaluation metric

Generally, evaluation metrics are employed to measure the quality of a classification model. There are many different types of evaluation metrics available to validate a classification model. In this work, we note, precision, sensitivity, specificity, f1-score and accuracy measures. Precision determines that out of all the depressed predicted, what percentage is truly depressed. Sensitivity also called as Recall tells us how many subjects in the diseased class are correctly classified. The ability of the method to reliably identify subjects without the ailment is known as specificity [32]. f1-score which is the harmonic mean of precision, and sensitivity [33]. Accuracy is determined by the capacity of the model to appropriately distinguish between depressed and healthy subjects. The precision, sensitivity, specificity, f1-score and accuracy measures are computed using Eqs. (4)–(9), respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (5)$$

Table 1
Classification Algorithms with their Parameters.

Classification_Algorithm	Parameters
KNN	n_neighbors = 7, weight = "uniform", algorithm = "kd_tree", leaf_size = 30, metric = "minkowski", n_jobs = 1
DT	criterion = 'gini', splitter = "best", max_depth = "None", min_samples_split = 2, max_features = None
XGBoost	n_estimators = 100, booster = "gbtree", learning_rate = 0.1, subsample = 1.0, criterion = "friedman_mse", min_samples_split = 2
SVM	kernel = "linear", degree = 3, gamma = "scale", shrinking = "True", tol = 1e-3, cache_size = 200, class_weight = "balanced"
LR	penalty = 12, dual = "false", tol = 1e-4, fit_intercept = "True", intercept_scaling = 1, class_weight = "balanced"
LDA	n_components = 1, solver = "svd", shrinkage = "None", priors = "None", tol = 1.0e-4, covariance_estimator = "None"

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (6)$$

$$FPR = 1 - \text{Specificity} \quad (7)$$

$$F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

where TP, TN, FP, and FN denote true-positive, true-negative, false-positive, and false-negative, respectively. In this study, the ROC curve is also used to graphically display a classifier's performance with varying threshold settings [34]. It is a probability curve and area under curve (auc) represents the degree or measure of separability to distinguish between classes. The ROC curve is plotted with TP rate (TPR) on the y-axis and FP rate (FPR) on the x-axis. The higher the auc, the better the model distinguish between subjects with and without the disease.

3.4. Dataset description

3.4.1. Dataset 1

The dataset used in the experiment is borrowed from a Multi-modal Open Dataset for Mental-disorder Analysis (MODMA) [35] which is openly available for depression analysis. The EEG signals were collected using a three-electrode Pervasive EEG collector at a frequency of 250 Hz. The EEG signals of 55 subjects were captured, out of which 23 were suffering from depression and the rest 33 were healthy subjects. The signals were collected from the pre-frontal part of the brain by placing three electrodes on the scalp at locations, namely Fp1, Fpz, and Fp3 based on a 10–20 system of electrode placement.

3.4.2. Dataset 2

The second dataset considered in this paper is taken from paper [36] which consists of EEG data signals of 33 subjects for 9 min at a resting state, out of which 15 subjects were depressed and 18 are healthy. The EEG data were collected using 20 electrodes at a frequency of 512 Hz. During the recording 0.1-Hz, high-pass filter, 100-Hz low-pass filter, and 50-Hz notch filter was employed to remove low-frequency noise and irrelevant baseline noise from the recorded data.

3.4.3. Dataset 3

The third dataset is also borrowed from MODMA that consists of an EEG recording of 53 subjects, out of which 24 subjects were depressed patients, out of which 13 were males and 11 were females, and 29 subjects were healthy, out of which 20 were males and 9 were females. The EEG signals were recorded using 128-channel HydroCel Geodesic Sensor Net (Electrical Geodesics Inc., Oregon Eugene, USA) and Net Station acquisition software at a sampling frequency of 256-Hz. To maintain good contact the impedance of the electrodes was checked and kept below 50 Kilo Ohms.

It should be noted that noise contained in EEG signals hinders increasing the classification accuracy of a machine learning algorithm to some extent, hence pre-processing step is required. Although many denoising techniques have been developed in previous studies [37,38], however the pre-processed versions of the raw datasets are already available publicly. Three such pre-processed datasets are considered for this study. Thus, no pre-processing step is required separately. The pre-processing steps widely used in the field of neuroscience permit to remove noise which is commonly known as artifact. There are mainly two types of artifacts: physiological and extra-physiological. Physiological are mainly related to movement artifacts and blinking artifacts. Extra-physiological are related mainly to the equipment and interference from the environment. However, how the pre-processing had been carried out is beyond the scope of the study. The interested readers are

referred to Refs. [35,36] to know more about the pre-processing step. In addition, the subjects were subjected to a clinical test assisted by professional psychologists to complete a preliminary judgment of depression by filling out a Patient Health Questionnaire (PHQ-9). The questionnaire consisted of 9 questions used for diagnosis, screening, and monitoring the severity of depression. It is based on Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria, with questions designed to detect symptoms of depression that have lasted at least two weeks. Based on it, a score known as the PHQ-9 score is assigned to each patient, which is employed to determine whether or not a subject is depressed. PHQ-9's sensitivity and specificity for identifying depressed patients have been reported to be 86.0% percent and 91.1% percent, respectively [39].

3.5. Fusion methods

Node2vec is used to extract relevant features from a dataset that will best represent it in a reduced dimension, as previously discussed. Three different types of fusion methods have been proposed at the graph, feature and decision levels described as follows:

3.5.1. Graph-level fusion

Graph-level fusion occurs immediately following the first step of graph construction as displayed in Fig. 3 (Appendix). In the first step, for each channel, a complete graph is created in which each node represents a subject in a dataset and the edge weight represents the euclidian distance between them. As a result, if the dataset contains n channels, namely $C_1, C_2, C_3 \dots C_n$, then n graphs that is $G_1, G_2, G_3 \dots G_n$ are generated each of which contains $\frac{p(p-1)}{2}$ edges where p is the total number of subjects in the dataset. The graph-level fusion method works by concatenating the graphs $G_1, G_2, G_3 \dots G_n$ and then constructing the final graph where each edge is computed using Eq. (10).

$$s_q = \sum_{i=1}^n w_{iq} \quad (10)$$

Here, q denotes the total number of edges present in the graph and s_q denotes a specific edge of the final graph. Consider if each of the complete graphs $G_1, G_2, G_3 \dots G_n$ consists of four nodes then a total of six edges are generated then the concatenated final graph will also contain four nodes and six edges, namely s_1, s_2, s_3, s_4, s_5 and s_6 whose value is computed using Eq. (10). It then generates a second order random walk for each node in the final graph, resulting in the creation of a corpus trained by the skip-gram model to obtain the embedding for each node in a two-dimensional space using the t-SNE algorithm [40]. These node embeddings of size size $n \times d$ act as feature vectors and are then passed to the classification algorithms to distinguish between depressed and healthy subjects.

3.5.2. Feature-level fusion

After the step of node embedding generation, the feature-level fusion is performed as shown in Fig. 4 (Appendix). After constructing n graphs, namely $G_1, G_2, G_3 \dots G_n$ from n different channels that is $C_1, C_2, C_3 \dots C_n$, random walks for each node are generated to form a corpus, which is then trained by the skip-gram model to generate node embeddings for each channel, namely $F_1, F_2, F_3 \dots F_n$, that act as feature vectors embedded in a two-dimensional space. The feature vectors generated by each channel are then combined to form the final set of feature vectors, which is called feature-level fusion. These feature vectors are of size $n \times d$ which are then fed into different classification algorithms, that categorize subjects as depressed or healthy. The size of d is taken to be 128.

3.5.3. Decision-level fusion

Another type of feature combination method used at the output level is decision level fusion as displayed in Fig. 3. After constructing n graphs, namely $G_1, G_2, G_3 \dots G_n$ from n different channels that is $C_1, C_2, C_3 \dots C_n$, random walks for each node are generated to form a corpus, which is

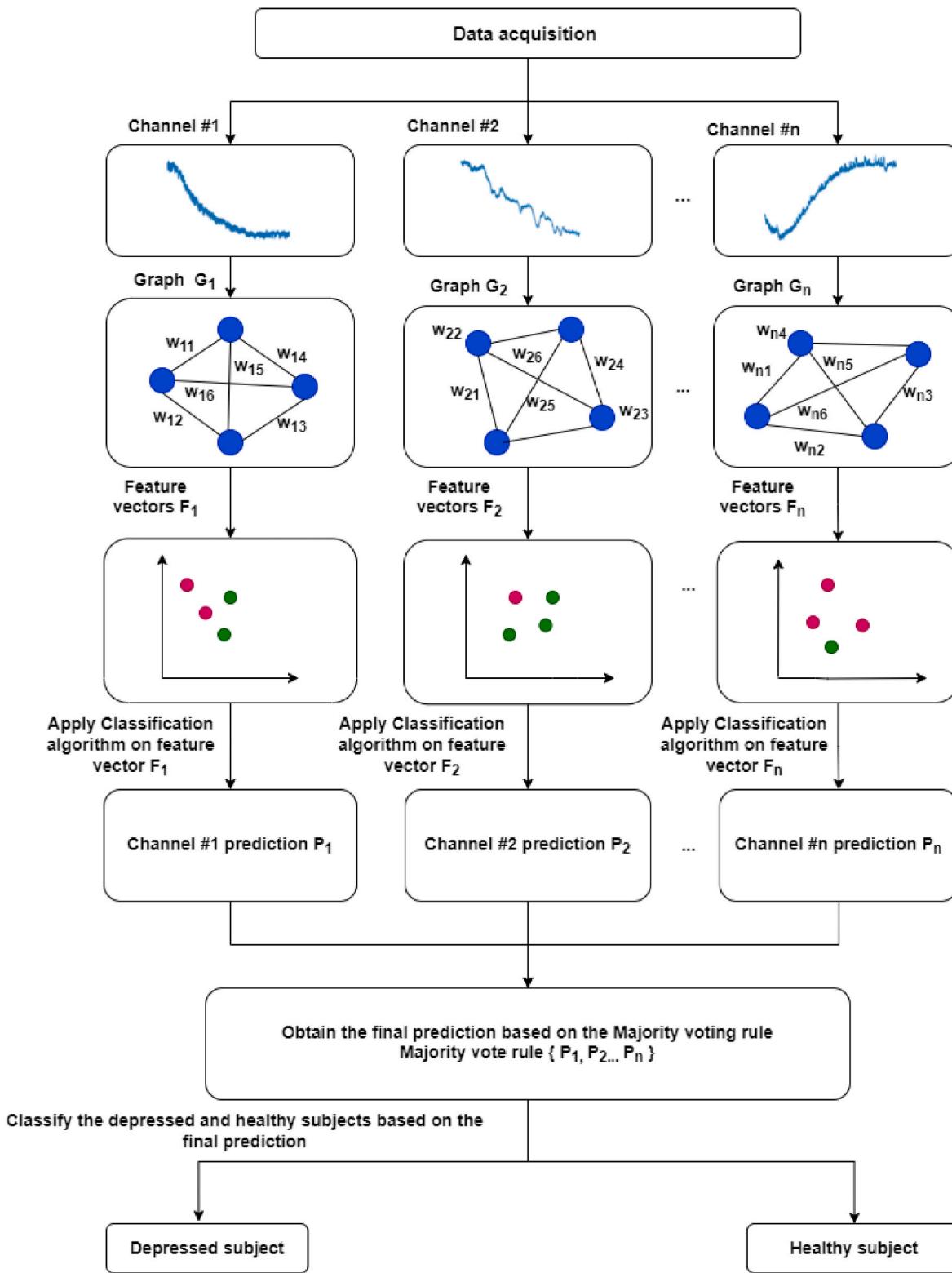


Fig. 3. Block diagram for decision-level fusion.

then trained by the skip-gram model to generate node embeddings for each channel, namely $F_1, F_2, F_3 \dots F_n$, that act as feature vectors embedded in a two-dimensional space. These feature vectors of size $n \times d$ are then passed to the classification algorithms to generate predictions $P_1, P_2, P_3 \dots P_n$ by dividing the feature vectors into training and testing. It compares the classifier's prediction decision in classifying the depressed and healthy groups for each channel and then makes the

decision to give the final predicted output based on the Majority function. It is the Boolean function that evaluates to false when half or more of the arguments are false and true otherwise, i.e. the function's value equals the value of the majority of the inputs. We can use the (real-valued) formula to represent depressed groups as 1 and healthy groups as 0 as shown in Eq. (11).

$$(p_1 \dots p_n) = Majority(p_1 \dots p_n) = \left\lfloor \frac{1}{2} + \frac{\left(\sum_{i=1}^n p_i \right) - \frac{1}{2}}{n} \right\rfloor \quad (11)$$

When the number of arguments n is even, the “1/2” in the formula breaks ties in favor of zeros otherwise the function breaks ties in favour of ones.

4. Experimental results and discussions

4.1. Experimental settings

The Keras framework and Google Colab platform have been used for the implementation of the experiment in this study. Python 3.8.1 language is employed for the implementation task using GPU RAM of 16 GB, System RAM of 15 GB, Intel(R) 2.30 GHz CPU, Tesla P100-PCIE Graphic Processor, and GDDR5X memory type.

4.2. Computational protocol

The proposed method's classification report is compared to three recent DL-based models and three machine learning approaches,

- **AchLSTM:** Automated depression detection using deep representation and sequence learning with EEG signals [41].
- **AchCNN:** Automated EEG-based screening of depression using deep CNN [42].
- **DeprNet:** A deep convolution neural network framework for detecting depression using EEG [36].
- **CaiH-KNN1:** A pervasive approach to EEG-based depression detection [26].
- **CaiH-KNN2:** Feature-Level fusion approaches for depression recognition based on multimodal EEG data [43].
- **CaiH-DBN** Pervasive EEG diagnosis of depression using Deep Belief Network with three-electrodes EEG collector [21].

We provide meaningful names because the scholars have not used any names to refer for their works. These state-of-the-art methods are implemented on the above-said three datasets. The above state-of-the-art methods are implemented as they are in the presented papers. The detailed description of these are beyond the scope of this study. The interested readers can refer to Refs. [41,42], [26,36,43], and [21] to know about these methods in detail.

4.3. Results

Node2vec is a simple, scalable, and successful technique for learning low-dimensional embeddings for nodes in a graph. The first step to implement the node2vec algorithm is the generation of a graph from the dataset, where the nodes of the graph represent the subjects present in the dataset and the edges between them represent the euclidian distance between them. Since the EEG data is in the form of a NumPy array hence it is feasible to determine the euclidian distance between them. After the formation of the complete graph, the algorithm creates sentences from this graph which is a list of node ids. Hence a collection of all sentences called the corpus is generated using four parameters: l , r , m , and o . The value of d here is taken as 128. The values of the parameters taken in the experiment are reflected in Table 2. Finally, the corpus is fed to the skip-

Table 2
Parameters table.

Parameters	Values
No of random walks (r)	10
length of a random walk (l)	100
Return Parameter (m)	0.5
In out Parameter (o)	2.0

gram model which gives a set of feature vectors in a two-dimensional plane where the blue point reflects healthy subjects and red points represents the depressed subjects.

Since dataset 1 includes 55 subjects hence, a complete graph with 55 nodes and 1540 edges is generated. Because the length of the random walks in our case is set to 10, 550 biased random walks are generated using the parameters m and n . The corpus of 550 random walks is then fed into the skip-gram model, which produces features in the form of node vectors of size 55×128 embedded in a two-dimensional plane.

In the case of dataset 2, the complete graph consists of 33 nodes and 551 edges. Because each node generates ten random walks, a total of 330 random walks are generated and fed to the skip-gram model, which generates feature vectors of size 33×128 embedded in a two-dimensional plane.

Similarly dataset 3 contains 53 subjects, with a complete graph of 53 nodes and 1431 edges. In this case, 530 biased random walks are generated and fed into the skip-gram model, which generates feature vectors of size 53×128 embedded in a two-dimensional plane.

4.3.1. Graph-level fusion results

As previously stated, graph-level fusion is carried out by concatenating the graphs obtained from the n channels and then using the Node2vec algorithm to obtain the feature vectors from the concatenated graph. The obtained feature vectors are then fed into the classification algorithms, which classifies the depressed and healthy subjects. As a result, three, twenty, and one twenty-eight graphs will be produced from datasets 1, 2 and 3 respectively. In case of dataset 1, the edges of three graphs are concatenated to form a final graph, on which the Node2vec algorithm is applied to generate a final set of features. Similarly, in the dataset 2, the edges of twenty graphs are concatenated to generate a feature set, and in the dataset 3, the edges of one twenty-eight graphs are concatenated to generate a feature set. The classification algorithms is then applied over the feature set to distinguish the healthy and depressed subjects.

In dataset 1, the result of the graph-level fusion is reported in Table 3. Among the five classification algorithms applied on the feature vectors which is obtained as a result of the Node2vec algorithm, the KNN classifier gives the highest accuracy of 0.785 in classifying the depressed and healthy subjects.

Apart from the accuracy metric, the auc metric is also taken into consideration to determine the performance of the proposed algorithm in distinguishing between depressed and healthy subjects. The KNN algorithm achieves the greatest auc of 0.875, which is the highest among all classifiers in dataset 1 as displayed in Fig. 4.

Similarly, for dataset 2, out of the five classifiers, the KNN algorithm gives the highest accuracy of 0.928 in classifying depressed and healthy subjects as noted in Table 4.

In dataset 2, the ROC curve is plotted and observed that the KNN algorithm provides the highest auc of 0.962 which is the highest among all the algorithms as shown in Fig. 5.

In the dataset 3, on classifying the feature vectors that are obtained by applying Node2vec algorithm on the graph generated as a result of graph-level fusion when passed to five classification algorithms, the KNN algorithm achieves the highest accuracy of 0.857 among all the different classifiers as noted in Table 5.

Table 3

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 1 for graph-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.743	0.754	0.782	0.748	0.785
SVM	0.612	0.632	0.629	0.621	0.647
LR	0.723	0.750	0.672	0.734	0.688
LDA	0.675	0.720	0.687	0.624	0.705
XGBoost	0.613	0.637	0.625	0.624	0.647
DT	0.698	0.688	0.673	0.692	0.667

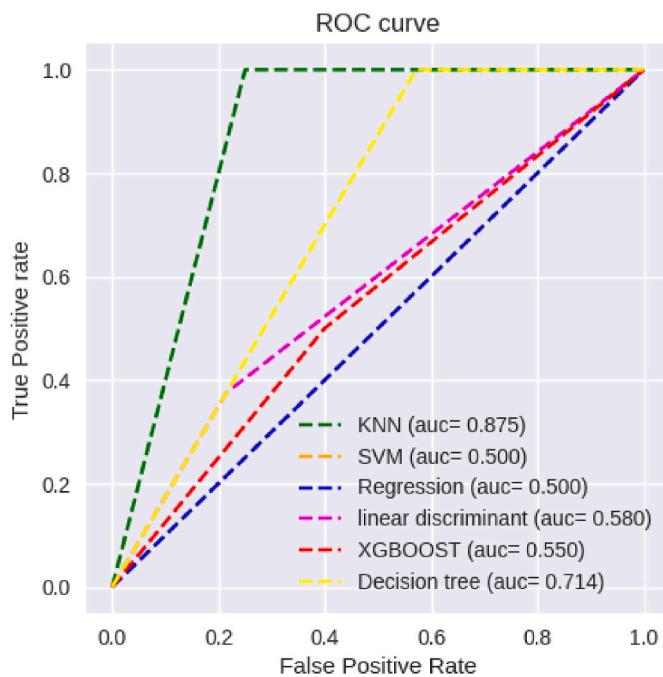


Fig. 4. ROC curve of dataset 1 for graph-level fusion.

Table 4

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 2 for graph-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.912	0.903	0.923	0.907	0.928
SVM	0.866	0.870	0.910	0.867	0.866
LR	0.857	0.860	0.847	0.858	0.837
LDA	0.713	0.723	0.730	0.717	0.733
XGBoost	0.767	0.790	0.730	0.778	0.823
DT	0.663	0.686	0.678	0.674	0.666

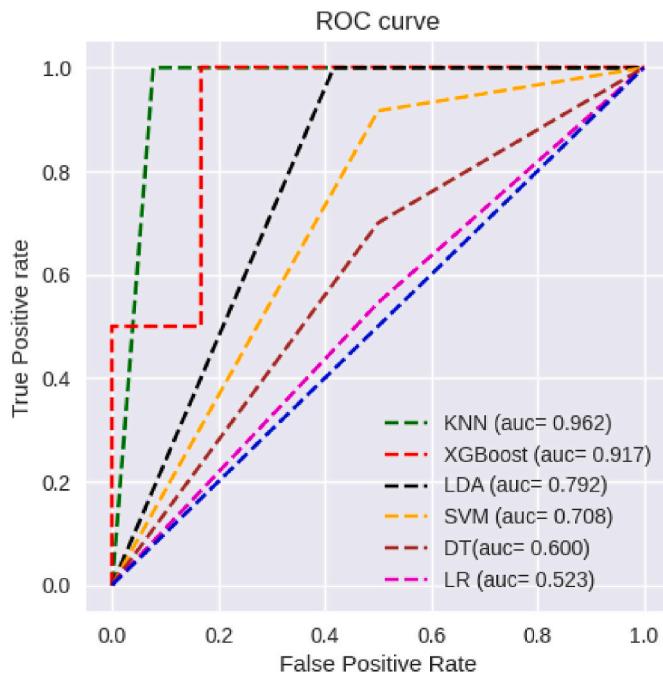


Fig. 5. ROC curve dataset 2 for graph-level fusion.

Table 5

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 3 for graph-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.822	0.862	0.832	0.841	0.857
SVM	0.817	0.823	0.802	0.819	0.833
LR	0.634	0.656	0.602	0.644	0.625
LDA	0.677	0.687	0.632	0.681	0.668
XGBoost	0.666	0.682	0.622	0.673	0.666
DT	0.666	0.672	0.612	0.668	0.625

In dataset 3, the ROC curve is plotted and observed that the KNN algorithm provides auc of 0.875 which is the highest among all the algorithms as displayed in Fig. 6.

4.3.2. Feature-level fusion results

As stated above in feature level fusion, the EEG signals extracted from different channels are passed individually to the Node2vec algorithm which provides the feature vectors of each channel embedded in a two-dimensional plane, which is then combined and passed to the classification algorithms. In the case of dataset 1, the feature vectors are obtained by combining the features of three channels are then fed into classification algorithms to differentiate between the depressed and healthy subjects. The results show that KNN outperforms the classification with the highest accuracy of 0.777 among the five algorithms used in the methodology as described in Table 6.

The ROC curve is plotted for feature-level fusion in the case of dataset 1 using different machine learning algorithms KNN algorithm provides the highest auc of 0.742 among all the other classification algorithms used in the proposed methodology as shown in Fig. 7.

For the dataset 2, the feature vectors are obtained by combining the feature vectors generated from the twenty channels when passed to the Node2vec algorithm which are then fed to the classification algorithms. The results show that among the five algorithms used in the methodology, KNN performs better with the highest accuracy of 0.909, as reported in Table 7.

The ROC curve plotted for feature level fusion in the case of the dataset 2 using classification algorithms provides the highest auc of 0.944 in the case of the LR algorithm among all the other algorithms as

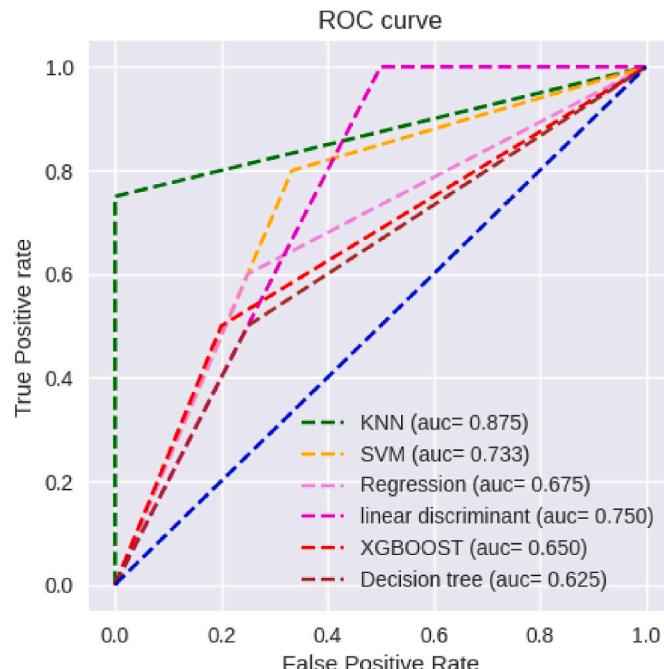
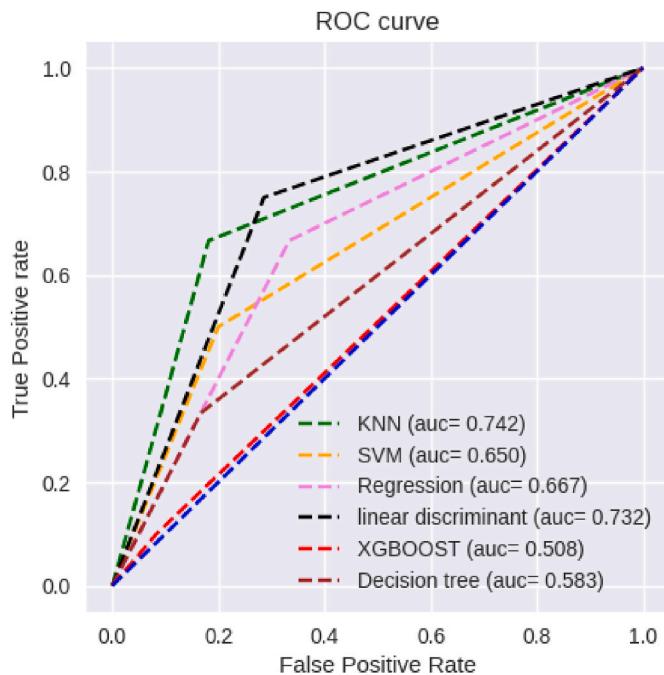


Fig. 6. ROC curve of dataset 3 for graph-level fusion.

Table 6

Precision, Sensitivity, Specificity, f1-scor, and Accuracy in dataset 1 for feature-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.776	0.783	0.792	0.779	0.777
SVM	0.622	0.603	0.643	0.612	0.666
LR	0.678	0.665	0.703	0.671	0.683
LDA	0.613	0.673	0.653	0.641	0.636
XGBoost	0.556	0.543	0.579	0.549	0.555
DT	0.524	0.533	0.563	0.528	0.575

**Fig. 7.** ROC curve of dataset 1 for feature-level fusion.**Table 7**

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 2 for feature-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.875	0.890	0.888	0.882	0.909
SVM	0.788	0.777	0.874	0.782	0.800
LR	0.822	0.833	0.876	0.827	0.908
LDA	0.755	0.733	0.789	0.743	0.802
XGBoost	0.843	0.822	0.868	0.832	0.900
DT	0.643	0.666	0.683	0.654	0.700

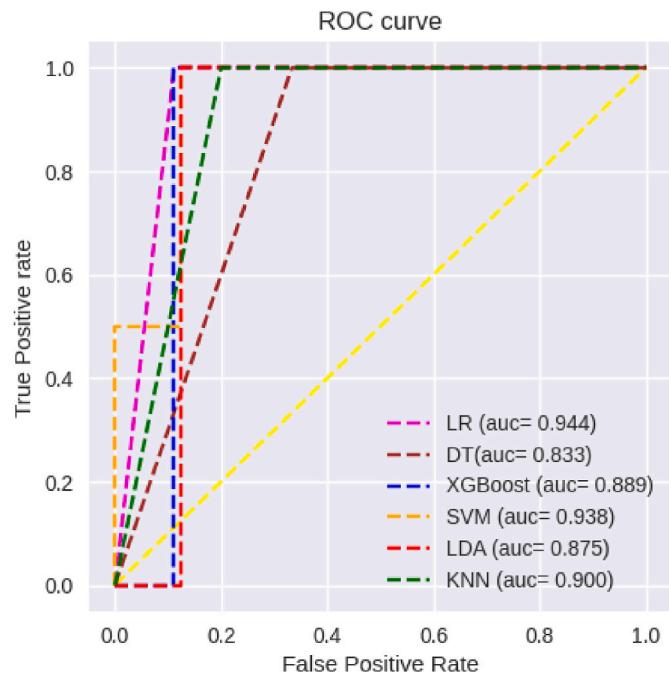
displayed in **Fig. 8**.

In the case of dataset 3, the feature vectors obtained by combining the feature vectors of one twenty-eight channels generated by passing the channels individually to the Node2vec algorithm are then passed to the classification algorithms to distinguish depressed and healthy subjects. Out of the five classification algorithms, the KNN algorithm achieves the highest accuracy of 0.833 as described in **Table 8**.

The ROC curve is plotted for feature-level fusion in the case of dataset 3. The KNN algorithm provides the highest auc of 0.944 among all the other algorithms, as shown in **Fig. 9**.

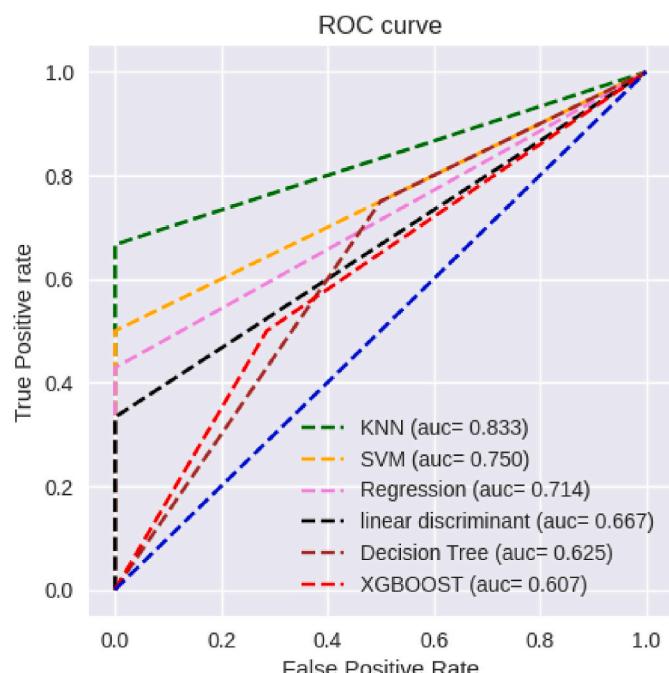
4.3.3. Decision-level fusion results

As previously stated, decision-level fusion is similar to feature-level fusion with the exception of one additional step. The channels of the EEG dataset were individually passed to the Node2vec algorithm in this

**Fig. 8.** ROC curve of dataset 2 for feature-level fusion.**Table 8**

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 3 for feature-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.799	0.811	0.822	0.804	0.833
SVM	0.711	0.732	0.722	0.721	0.752
LR	0.698	0.703	0.702	0.700	0.714
LDA	0.673	0.693	0.687	0.682	0.704
XGBoost	0.587	0.607	0.623	0.596	0.636
DT	0.579	0.542	0.612	0.559	0.606

**Fig. 9.** ROC curve of dataset 3 for feature-level fusion.

fusion method, and the feature vectors obtained are passed to the classifiers for classification. The predictions obtained from the classifiers for each channel are then analyzed, and a final prediction based on the majority voting rule is obtained. The predictions obtained by the three channels when the feature vectors of each channel are passed to the classification algorithms in the case of dataset 1 are then analyzed and a final prediction decision is made. The KNN algorithm achieves the highest accuracy of 0.823 in classifying depressed and healthy groups using decision-level fusion among all five classification algorithms used in this methodology, as reported in [Table 9](#).

The ROC curve plotted in the case of the dataset 1 gave the highest auc of 0.875 with the KNN algorithm when compared to the other algorithms as displayed in [Fig. 10](#).

In the case of dataset 2, the predictions that are obtained by the twenty channels when the feature vectors of each channel are passed to the classification algorithm are analyzed and a final prediction decision is taken. The highest accuracy of 0.933 is obtained with the KNN using the decision-level fusion among all the other five classification algorithm used in this methodology as noted in [Table 10](#).

In the case of dataset 2, the ROC curve plotted gives the highest auc of 0.964 in the case of the KNN algorithm when compared to the other algorithms as displayed in [Fig. 11](#).

The predictions obtained in case of dataset 3 when the feature vectors of each channel are passed to the classification algorithms were analyzed and a final prediction decision is made based on Majority voting rule. The KNN algorithm achieves the highest accuracy of 0.888 among all five classification algorithms used in this methodology, as described in [Table 11](#).

The ROC curve plotted in the case dataset 3 gives the highest auc of 0.875 with the KNN algorithm when compared to the other algorithms as shown in [Fig. 12](#).

4.4. Comparative study

The proposed method is compared three deep learning methods and three handcrafted feature methods on three datasets that are collected from three different sources. The aim of this work is to check how good the proposed method is over state-of-the-art methods. A comparative analysis to benchmark state-of-the-art methods for depression detection using EEG signals with respect to validation metrics, such as precision, recall, F1-score, and accuracy, is performed. However, the meta-data of these datasets, for example, the sample rate, data size, etc, are not considered, which deserve further study. The comparative study demonstrates that the proposed method outperforms all other methods in terms of validation metrics in detecting depressed and healthy subjects as described in [Table 12](#).

4.5. Empirical study

The Node2vec algorithm generates feature vectors by performing biased random walks using four parameters m , o , l , and r . The value of these parameters is decided to be optimal by an experimental study performed on the datasets taking the different values of the parameters.

Table 9

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 1 for decision-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.777	0.818	0.789	0.796	0.823
SVM	0.727	0.723	0.666	0.724	0.714
LR	0.634	0.689	0.646	0.660	0.666
LDA	0.724	0.714	0.750	0.718	0.727
XGBoost	0.513	0.533	0.562	0.522	0.545
DT	0.688	0.673	0.697	0.680	0.683

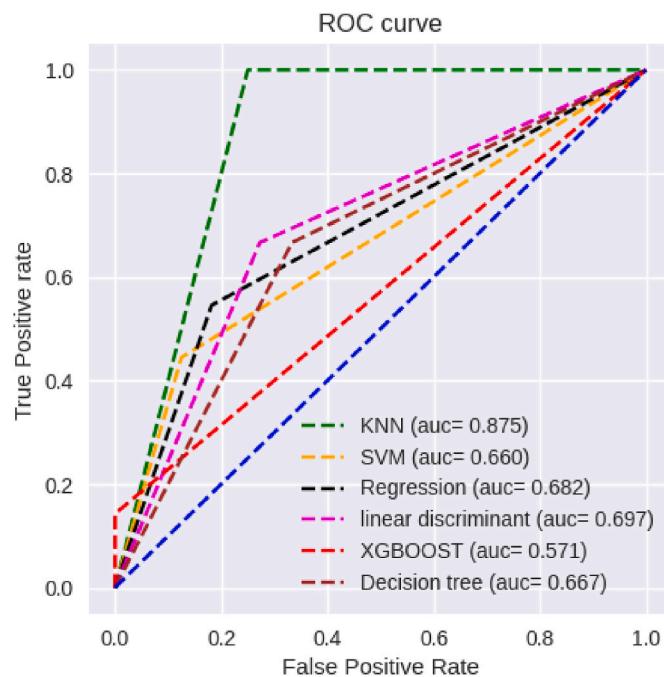


Fig. 10. ROC curve of dataset 1 for decision-level fusion.

Table 10

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 2 for decision-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.895	0.915	0.923	0.904	0.933
SVM	0.865	0.859	0.908	0.861	0.900
LR	0.877	0.899	0.843	0.875	0.916
LDA	0.832	0.865	0.900	0.848	0.923
XGBoost	0.813	0.842	0.823	0.836	0.866
DT	0.708	0.710	0.702	0.708	0.733

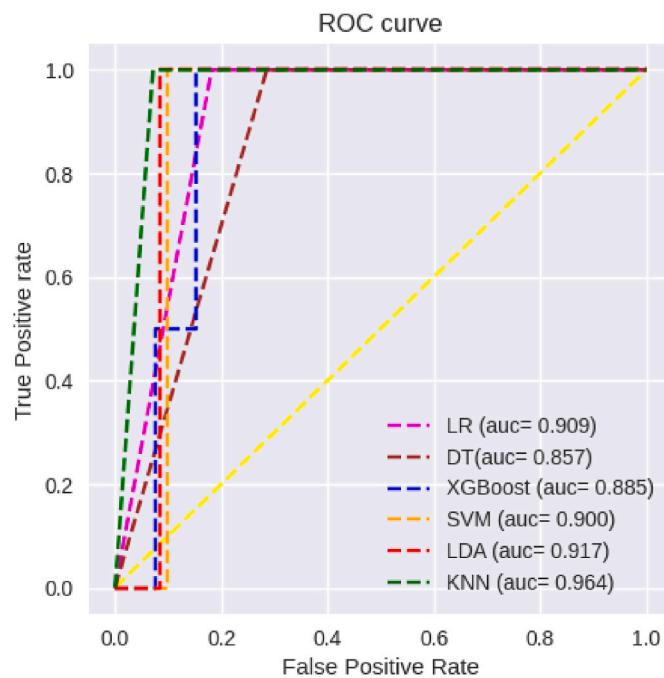
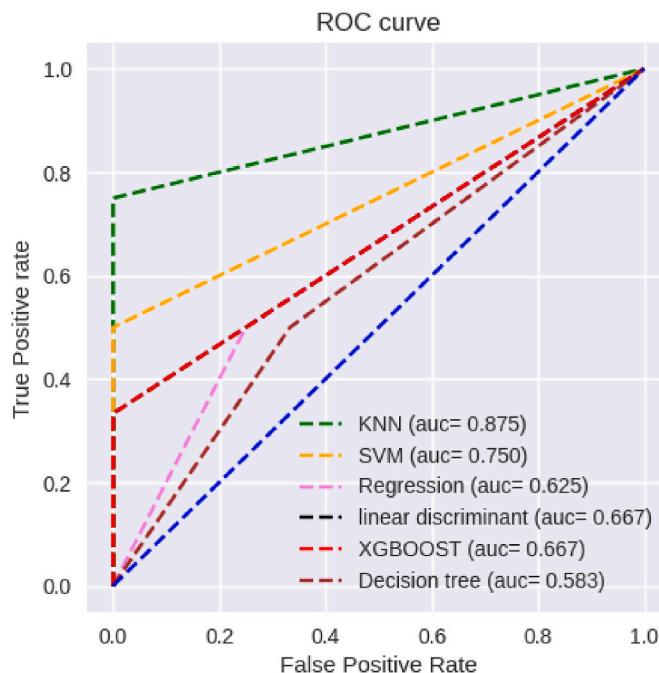


Fig. 11. ROC curve of dataset 2 for decision-level fusion.

Table 11

Precision, Sensitivity, Specificity, F1-score, and Accuracy in dataset 3 for decision-level fusion.

Algorithm	Precision	Sensitivity	Specificity	F1-score	Accuracy
KNN	0.852	0.872	0.863	0.861	0.888
SVM	0.743	0.723	0.733	0.737	0.750
LR	0.788	0.803	0.792	0.795	0.818
LDA	0.743	0.827	0.803	0.782	0.833
XGBoost	0.666	0.653	0.627	0.659	0.666
DT	0.687	0.702	0.692	0.694	0.714

**Fig. 12.** ROC curve of dataset 3 for decision-level fusion.

4.5.1. Varying r with constants m, o, and l

The values of m, o, and l are kept constant as reported in [Table 2](#), but different values of r are used, and the change in accuracy is observed. The values taken are 5, 10, 15, 20, and 25. It is discovered that the highest accuracy is obtained when the value of r is set to 10. The experiment is carried out on all three datasets, and the variation in accuracy with change in r value is shown in [Fig. 13](#).

4.5.2. Varying l with constants m, o, and r

To observe the change in accuracy, the values of m, o, and r are kept constant as specified in [Table 2](#), and different values of l are used. The values taken are 50, 100, 150, 200, 250, and 300. During the experiment, it is discovered that the case of 100 yielded the highest accuracy. The experiment is carried out on all three datasets, and the variation in

accuracy with change in l value is demonstrated in [Fig. 14](#).

4.5.3. Varying m with constants l, o, and r

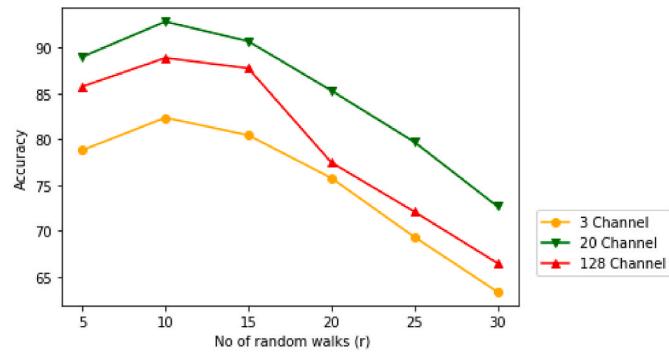
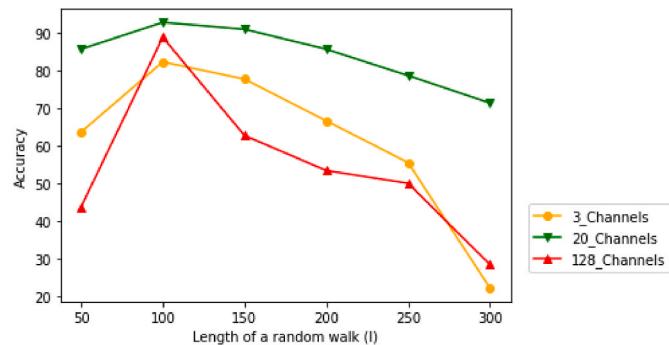
The values of l, o, and r remains constant as noted in [Table 2](#), while different values of m are used to observe the change in accuracy. The values chosen are 0.1, 0.3, 0.5, 0.7, 0.9, and 1.0. While experimenting, it is discovered that the case of 0.5 yielded the highest accuracy. The experiment is carried out on all three datasets, and the variation in accuracy with change in m value is shown in [Fig. 15](#).

4.5.4. Varying o with constants r, l, and m

The values of r, l, and m are held constant as specified in [Table 2](#), and the accuracy is measured for a variety of o values. The values of n during the experiment are 0.5, 1.0, 1.5, 2.0, and 2.5. It is discovered that when the value of o are set to 2.0, the highest accuracy is obtained. The experiment is carried out across all three datasets, and the variation in accuracy with change in o value is shown in [Fig. 16](#).

4.6. Discussion

We present a discussion on each of the methods implemented and

**Fig. 13.** Change in Accuracy with variation in values of r.**Fig. 14.** Change in Accuracy with variation in values of l.**Table 12**

Comparative study: The classification report of AchCNN, AchLSTM, DeprNet, CaiHH-KNN1, CaiH-KNN2, CaiH-DBN and proposed method for dataset 1, dataset 2 and dataset 3.

Method	Dataset 1				Dataset 2				Dataset 3			
	Precision	Recall	f1-score	Accuracy	Precision	Recall	f1-score	Accuracy	Precision	Recall	f1-score	Accuracy
AchCNN [42]	0.564	0.612	0.587	0.635	0.581	0.639	0.610	0.681	0.702	0.713	0.707	0.732
AchLSTM [41]	0.607	0.643	0.624	0.678	0.613	0.913	0.734	0.744	0.744	0.733	0.738	0.753
DeprNet [36]	0.651	0.688	0.668	0.702	0.919	0.887	0.895	0.914	0.775	0.723	0.748	0.782
\CaiH-KNN1 [26]	0.666	0.633	0.649	0.684	0.527	0.655	0.588	0.723	0.688	0.697	0.692	0.703
CaiH-KNN2 [43]	0.668	0.704	0.685	0.713	0.667	0.735	0.699	0.724	0.688	0.701	0.694	0.723
\CaiH-DBN\} [21]	0.622	0.597	0.609	0.635	0.617	0.680	0.647	0.671	0.652	0.643	0.647	0.662
Proposed Method	0.789	0.818	0.803	0.823	0.888	0.915	0.901	0.933	0.866	0.872	0.871	0.888

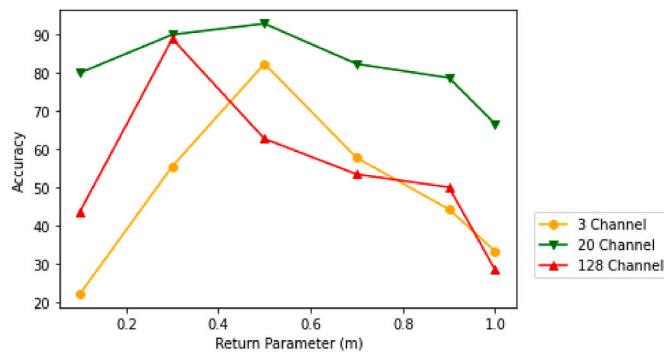


Fig. 15. Change in Accuracy with variation in values of m.

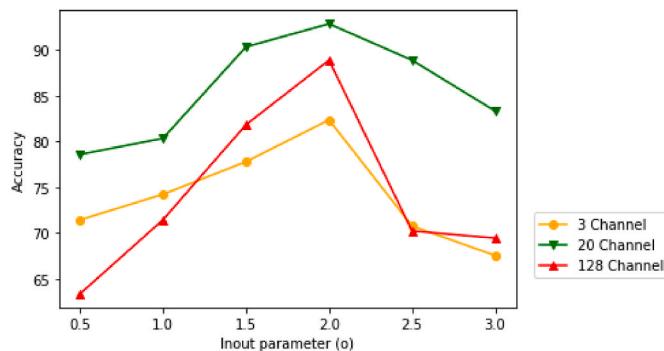


Fig. 16. Change in Accuracy with variation in values of o.

analyzed in this paper. Relying solely on traditional approaches, such as questionnaires, might fail to detect depression. With the technological advances in brain signal analysis, human brain activities in the form of non-invasive EEG recordings can produce more reliable results. In this work, three fusion methods are proposed at three different levels. However, all the three above-said methods exploit Node2vec algorithmic framework. This study evaluates the performance of the proposed methods on the three publicly available pre-processed datasets consisting of 3, 20, and 128 channels, respectively, using different evaluation metrics. Out of them, the performance of the proposed method at the decision level is superior since it tries to combine the decisions of several channels to produce a single final decision using majority voting technique. The decision of each channel is the classification conducted on the test dataset. However, the contribution of the individual channels of these datasets towards the performance of the method is not analyzed and will be carried out in near future.

5. Conclusion

The study successfully uses the graph representation learning approach for automatically extracting features from each channel and applies three different types of fusion, namely a graph-level fusion, feature-level fusion, and decision-level fusion for analyzing EEG data and classifying healthy and depressed subjects. The proposed method can successfully distinguish the healthy and depressed subjects with the highest accuracy of 0.933, sensitivity of 0.916, specificity of 0.923, the precision of 0.895, f1-score of 0.904, and an auc of 0.966 with the KNN Algorithm in case of decision-level fusion that outperforms the other state-of-art methods. The proposed method's accuracy is promising, but there is still potential for improvement. The method works by first constructing a graph from the dataset, in which the nodes represent the subjects in the dataset and where the edge weights obtained using the Euclidean distance reflect their relationship. The Node2vec algorithm is applied over the graph to create distinct features in the form of node

embeddings, which are then fed to classification algorithms to classify the depressed and healthy subjects. Our next steps will focus on creating a sparse graph rather than a complete graph, which will aid in the extraction of more different characteristics and improve the model's accuracy. Furthermore, due to the very complicated and informative graph structure, machine learning on graphs is known to be a difficult task. As a result, future research will concentrate on using graph convolutional networks (GCNs) to work directly on graphs, leveraging structural information and producing useful feature representations for nodes in the graph, which will help to improve the method's accuracy in detecting depressed subjects.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgment

This work is partially supported by the project IT4Neuro (degeneration), reg. nr. CZ.02.1.01/0.0/0.0/18 069/0010054 and by the project "Smart Solutions in Ubiquitous Computing Environments", Grant Agency of Excellence, University of Hradec Kralove, Faculty of Informatics and Management, Czech Republic (under ID: UHK-FIM-GE-2022).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.combiomed.2022.105420>.

References

- [1] W.H. Organization, et al., Depression and other common mental disorders: global health estimates, Tech. rep. (2017) 1–24.
- [2] A. Shoeibi, N. Ghassemi, R. Alizadehsani, M. Rouhani, H. Hosseini-Nejad, A. Khosravi, M. Panahiazar, S. Nahavandi, A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in eeg signals, Expert Systems with Applications 163 (2021) 1–16.
- [3] S.K. Khare, V. Bajaj, U.R. Acharya, Detection of Parkinson's disease using automated tunable q wavelet transform technique with eeg signals, Biocybernetics and Biomedical Engineering 41 (2021) 679–689.
- [4] M. Sharma, U.R. Acharya, Automated Detection of Schizophrenia Using Optimal Wavelet-Based Norm Features Extracted from Single-Channel Eeg, Cognitive Neurodynamics, 2021, pp. 1–14.
- [5] A. Asif, M. Majid, S.M. Anwar, Human stress classification using eeg signals in response to music tracks, Computers in biology and medicine 107 (2019) 182–196.
- [6] M. Sharma, J. Tiwari, U.R. Acharya, Automatic sleep-stage scoring in healthy and sleep disorder patients using optimal wavelet filter bank technique with eeg signals, International Journal of Environmental Research and Public Health 18 (2021) 1–29.
- [7] M.M. Rahman, A.K. Sarkar, M.A. Hossain, M.S. Hossain, M.R. Islam, M.B. Hossain, J.M. Quinn, M.A. Moni, Recognition of human emotions using eeg signals: a review, Computers in Biology and Medicine 136 (2021) 104696.
- [8] H. Hinrikus, A. Suhrova, M. Bachmann, K. Aadamssoo, Ü. Vöhma, J. Lass, V. Tuulik, Electroencephalographic spectral asymmetry index for detection of depression, Medical & biological engineering & computing 47 (12) (2009) 1291–1299.
- [9] V.A. Grin-Yatsenko, I. Baas, V.A. Ponomarev, J.D. Kropotov, Independent component approach to the analysis of eeg recordings at early stages of depressive disorders, Clinical Neurophysiology 121 (2010) 281–289.
- [10] J.L. Stewart, J.A. Coan, D.N. Towers, J.J. Allen, Resting and task-elicited prefrontal eeg alpha asymmetry in depression: support for the capability model, Psychophysiology 51 (2014) 446–455.
- [11] G.E. Bruder, J.W. Stewart, D. Hellerstein, J.E. Alvarenga, D. Alschuler, P. J. McGrath, Abnormal functional brain asymmetry in depression: evidence of biologic commonality between major depression and dysthymia, Psychiatry research 196 (2012) 250–254.
- [12] W. Heller, M.A. Etienne, G.A. Miller, Patterns of perceptual asymmetry in depression and anxiety: implications for neuropsychological models of emotion and psychopathology, Journal of abnormal psychology 104 (1995) 327–333.
- [13] S.D. Puthankattil, P.K. Joseph, Classification of eeg signals in normal and depression conditions by ann using rwe and signal entropy, Journal of Mechanics in Medicine and biology 12 (2012) 1240019.
- [14] A. Gaouda, M. Salama, M. Sultan, A. Chikhani, Power quality detection and classification using wavelet-multiresolution signal decomposition, IEEE Transactions on power delivery 14 (1999) 1469–1476.

- [15] M. Ahmadlou, H. Adeli, A. Adeli, Fractality analysis of frontal brain in major depressive disorder, *International Journal of Psychophysiology* 85 (2012) 206–211.
- [16] M. Bachmann, J. Lass, A. Suhhova, H. Hinrikus, Spectral asymmetry and higuchi's fractal dimension measures of depression electroencephalogram, *Computational and mathematical methods in medicine* 2013 (2013) 1–9.
- [17] T.Q.D. Khoa, V.Q. Ha, V.V. Toi, Higuchi fractal properties of onset epilepsy electroencephalogram, *Computational and mathematical methods in medicine* 2012 (2012) 1–7.
- [18] B. Hosseiniyfard, M.H. Moradi, R. Rostami, Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from eeg signal, *Computer methods and programs in biomedicine* 109 (2013) 339–345.
- [19] O. Faust, P.C.A. Ang, S.D. Puthankattil, P.K. Joseph, Depression diagnosis support system based on eeg signal entropies, *Journal of mechanics in medicine and biology* 14 (2014) 1450035.
- [20] U.R. Acharya, V.K. Sudarshan, H. Adeli, J. Santhosh, J.E. Koh, S.D. Puthankattil, A. Adeli, A novel depression diagnosis index using nonlinear features in eeg signals, *European neurology* 74 (2015) 79–83.
- [21] H. Cai, X. Sha, X. Han, S. Wei, B. Hu, Pervasive eeg diagnosis of depression using deep belief network with three-electrodes eeg collector, in: 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2016, pp. 1239–1246.
- [22] W. Mumtaz, L. Xia, S.S.A. Ali, M.A.M. Yasin, M. Hussain, A.S. Malik, Electroencephalogram (eeg)-based computer-aided technique to diagnose major depressive disorder (mdd), *Biomedical Signal Processing and Control* 31 (2017) 108–115.
- [23] H. Liu, H. Motoda, *Computational Methods of Feature Selection*, CRC Press, 2007, pp. 1–440.
- [24] S.-C. Liao, C.-T. Wu, H.-C. Huang, W.-T. Cheng, Y.-H. Liu, Major depression detection from eeg signals using kernel eigen-filter-bank common spatial patterns, *Sensors* 17 (2017) 1385.
- [25] M. Sharma, P. Achuth, D. Deb, S.D. Puthankattil, U.R. Acharya, An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with eeg signals, *Cognitive Systems Research* 52 (2018) 508–520.
- [26] H. Cai, J. Han, Y. Chen, X. Sha, Z. Wang, B. Hu, J. Yang, L. Feng, Z. Ding, Y. Chen, et al., A pervasive approach to eeg-based depression detection, *Complexity* 2018 (2018), 5238028.
- [27] S. Byun, A.Y. Kim, E.H. Jang, S. Kim, K.W. Choi, H.Y. Yu, H.J. Jeon, Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol, *Computers in biology and medicine* 112 (2019) 103381.
- [28] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, B. Hu, Feature-level fusion approaches based on multimodal eeg data for depression recognition, *Information Fusion* 59 (2020) 127–138.
- [29] A. Grover, J. Leskovec, node2vec: scalable feature learning for networks, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2016, pp. 855–864.
- [30] Y. Wu, Y. Bian, X. Zhang, Remember where you came from: on the second-order random walk based proximity measures, *Proceedings of the VLDB Endowment* 10 (2016) 13–24.
- [31] S. Chauhan, M. Singh, A.K. Aggarwal, Data science and data analytics: Artificial intelligence and machine learning integrated based approach, *Data Science and Data Analytics: Opportunities and Challenges* (2021) 1.
- [32] A. Kaur, A.P.S. Chauhan, A.K. Aggarwal, An automated slice sorting technique for multi-slice computed tomography liver cancer images using convolutional network, *Expert Systems with Applications* 186 (2021) 115686.
- [33] A. Kaur, A.P.S. Chauhan, A.K. Aggarwal, Machine learning based comparative analysis of methods for enhancer prediction in genomic data, in: 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), IEEE, 2019, pp. 142–145.
- [34] A.P. Bradley, The use of the area under the roc curve in the evaluation of machine learning algorithms, *Pattern recognition* 30 (1997) 1145–1159.
- [35] H. Cai, Y. Gao, S. Sun, N. Li, F. Tian, H. Xiao, J. Li, Z. Yang, X. Li, Q. Zhao, et al., Modma Dataset: a Multi-Modal Open Dataset for Mental-Disorder Analysis, 09283, 2002. arXiv preprint arXiv.
- [36] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, O. Krejcar, Deprnet: a deep convolution neural network framework for detecting depression using eeg, *IEEE Transactions on Instrumentation and Measurement* 70 (2021) 1–13.
- [37] R. Thukral, A. Arora, A. Kumar, et al., Denoising of thermal images using deep neural network, in: *Proceedings of International Conference on Recent Trends in Computing*, Springer, 2022, pp. 827–833.
- [38] R. Thukral, A. Kumar, A. Arora, et al., Effect of different thresholding techniques for denoising of emg signals by using different wavelets, in: 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), IEEE, 2019, pp. 161–165.
- [39] S.-I. Liu, Z.-T. Yeh, H.-C. Huang, F.-J. Sun, J.-J. Tjung, L.-C. Hwang, Y.-H. Shih, A. W.-C. Yeh, Validation of patient health questionnaire for depression screening among primary care patients in taiwan, *Comprehensive psychiatry* 52 (2011) 96–101.
- [40] L. Van der Maaten, G. Hinton, Visualizing data using t-sne, *Journal of machine learning research* 9 (2008) 1–27.
- [41] B. Ay, O. Yildirim, M. Talo, U.B. Baloglu, G. Aydin, S.D. Puthankattil, U.R. Acharya, Automated depression detection using deep representation and sequence learning with eeg signals, *Journal of medical systems* 43 (7) (2019) 1–12.
- [42] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, D.P. Subha, Automated eeg-based screening of depression using deep convolutional neural network, *Computer methods and programs in biomedicine* 161 (2018) 103–113.
- [43] S. Zheng, C. Lei, T. Wang, C. Wu, J. Sun, H. Peng, Feature-level fusion for depression recognition based on fmrs data, in: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2020, pp. 2906–2913.