Applied Intelligence (2021) 51:6449–6466
https://doi.org/10.1007/s10489-021-02426-y

Automated major depressive disorder detection using melamine pattern with EEG signals

Emrah Aydemir[1] · Turker Tuncer[2] · Sengul Dogan[2] · Raj Gururajan[3] · U. Rajendra Acharya[4,5,6]

Accepted: 8 April 2021 / Published online: 28 April 2021

Abstract
Major depressive disorder (MDD) is one of the most common modern ailments affected huge population throughout the world. The electroencephalogram (EEG) signal is widely used to screen the MDD. The manual diagnosis of MDD using EEG is time consuming, subjective and may cause human errors. Therefore, nowadays various automated systems have been developed to diagnose MDD accurately and rapidly. In this work, we have proposed a novel automated MDD detection system using EEG signals. Our proposed model has *three* steps: (i) Melamine pattern and discrete wavelet transform (DWT)- based multileveled feature generation, (ii) selection of most relevant features using neighborhood component analysis (NCA) and (iii) classification using support vector machine (SVM) and k nearest neighbor (kNN) classifiers. The novelty of this work is the application of melamine pattern. The molecular structure of melamine (also named chemistry spider- ChemSpider) is used to generate 1536 features. Also, various statistical features are extracted from DWT coefficients. The NCA is used to select the most relevant features and these selected features are classified using SVM and kNN classifiers. The presented model attained greater than 95% accuracies using all channels with quadratic SVM classifier. Our results obtained highest classification accuracy of 99.11% and 99.05% using Weighted kNN and Quadratic SVM respectively using A2A1 EEG channel. We have developed the automated depression model using a big dataset and yielded high classification accuracies. These results indicate that our presented model can be used in mental health clinics to confirm the manual diagnosis of psychiatrists.

Keywords Melamine pattern · Statistical feature generation · Major depression detection · NCA selector · EEG signal processing

1 Introduction

Sadness is one of the moods that any person can experience and is not a sign of disorder if it lasts for a short duration. However, prolonged sadness is one of the most obvious symptoms of depression. Sadness that lasts for at least two weeks is referred to as major depressive disorder (MDD) [1–3]. About 300 million people worldwide has suffer from MDD [4]. In addition, the number of female MDD patients is approximately 1.5–2 times higher than male MDD patients [5]. The widely seen symptoms of MDD are low self-esteem, sadness, tearfulness, burst of rage, suicidal tendency and hallucinations. Therefore, it will affect their social lives and hence, they cannot interact with other people easily [6, 7].

MDD is a mental disorder and it should be diagnosed by specialist physicians. It can be treated with therapy and

✉ U. Rajendra Acharya
aru@np.edu.sg

Emrah Aydemir
emrahaydemir@sakarya.edu.tr

Turker Tuncer
turkertuncer@firat.edu.tr

Sengul Dogan
sdogan@firat.edu.tr

Raj Gururajan
Raj.Gururajan@usq.edu.au

[1] Department of Management Information, College of Management, Sakarya University, Sakarya, Turkey

[2] Department of Digital Forensics Engineering, College of Technology, Firat University, Elazig, Turkey

[3] School of Management and Enterprise, University of Southern Queensland, Toowoomba, Australia

[4] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489, Singapore

[5] Department of Biomedical Engineering, School of Science and Technology, SUSS University, Singapore, Singapore

[6] Department of Biomedical Informatics and Medical Engineering, Asia University, Taichung, Taiwan

medication (drug treatment). The commonly used therapies are cognitive behavioral therapy, electroconvulsive therapy, and interpersonal therapy. Drug treatment is used to prevent unwanted behaviors in advanced MDD cases [8, 9].

Usually it is very difficult to detect MDD manually in our daily life and also patients suffering from this disorder are reluctant to seek treatment due to social stigma. Therefore, early diagnosis and treatment of MDD is most important. Hamilton Depression Rating Scale and Beck Depression Inventory have been developed as a manual for the diagnosis of MDD and generally professionals use questionnaires to patients for diagnosis [6]. There is possibility of subjectivity in this and also patients may lie. Therefore, the most reliable way of MDD diagnosis is electroencephalogram (EEG) based diagnosis. The EEG signals consist of information about the neuronal activities of the brain. These EEG signals have the signatures of activities of the brain [10]. Using novel features extraction techniques the hidden information about the MDD can be obtained from the EEG signals [11]. However, it difficult to obtain salient information from the EEG signals as they are nonlinear and non-stationary in nature. Therefore, nonlinear features extraction methods coupled with machine learning techniques need to be employed to obtain high detection performances [12, 13]. Many machine learning techniques have widely used for automated detection of diseases and assist the clinicians [14–17].

In this study, we have developed MDD detection from EEG signals using machine learning techniques. We have used 34 MDD and 30 healthy subjects to develop the automated system. A new melamine pattern is proposed to generate the features from the EEG signals. These features are used to for the detection of MDD.

1.1 Literature review

In this section, depression studies conducted using speech, voice and facial biomarkers is provided Table 1. This table clearly shows that, authors did not obtain high detection accuracy using speech, voice and facial biomarkers as they are unable to extract unique features for normal and depression patients. Hence, many authors have used EEG signals in their studies and obtained higher detection performances [11].

1.2 Motivation and our model

Table 1 shows the summary of studies conducted to develop the MDD detection systems using speech, voice and facial biomarkers. However, authors did not obtain higher detection accuracies. Therefore, in this work, we are proposing to develop an automated MDD detection model using EEG signals. In this work, we have used 20 channel EEG signals with 7339 EEG signals of 10 s duration in each channel. For this big dataset, many deep learning models have been used [36–39]. Generally, deep models have high computational complexities and take longer time to train. For instance, ten-millions parameters should be set in the convolutional neural networks such as residual network, dense network, and Google network [40–42]. Therefore, in this work, we have used hand-crafted features based classification model. This model includes feature generation, feature selection and classification phases. Our proposed melamine pattern and NCA- based depression detection model is shown in Fig. 1.

The novelty of this work is the use of melamine pattern (it is in the feature generation box). It uses molecular structure (molecular graph) of the melamine and used molecular graph called ChemSpider. Our aim is to show the superiority of molecular structured feature generation model. To gain high classification accuracies, features are extracted from both low and high frequency components of discrete wavelet transform (DWT) [43]. The neighborhood component analysis (NCA) is employed on the generated features and 256 features [44] are selected. These features are fed to support vector machine (SVM) and k nearest neighbor (kNN) [45, 46] for automated classification. Our classifiers are developed using hold-out validation (80:20) strategy.

1.3 Key contributions

Key contributions of the proposed melamine pattern and NCA- based model are given below:

- A new molecular structure based feature generation model is presented.
- Proposed model is accurate and robust as we have obtained classification accuracy of more than 95% for all channels.

2 Materials

Mumtaz [47] collected EEG dataset in 2016 to develop an automated MDD. The EEG signals were collected from 64 subjects between the ages of 12 and 77 years (average age = 20.54 years). Among the collected 64 subjects (30 healthy and 34 MDD patients), 40 of them were men and 24 were women. The EEG signals were collected between 23.06.2011 and 30.06.2013. In this dataset, the EEG signals were sampled at 256 Hz and segmented into 10-s length. Therefore, the length of each EEG signal is 2560 samples. This corpus includes 20 channels with 7339 (3893 normal and 3446 MDD) EEG signals in each channel. The twenty channels used are A2-A1, C3, C4, Cz, F3, F4, F7, F8, Fp1, Fp2, Fz, O1, O2, P3, P4, Pz, T3, T4, T5 and T6. The sample normal and MDD EEG signals are shown in Fig. 2.

Table 1 Summary of studies done detecting depression using speech, voice and facial biomarkers

Study	Method	Material	Results
Jiang et al. [18]	Task-related common spatial patter	EEG	Accuracy: 84.00% for positive stimuli 85.70% for negative stimuli
Sharma et al. [19]	Depression hybrid neural network	EEG	Accuracy: 99.10%
Akbari et al. [20]	Empirical wavelet transform, Centered correntropy,	EEG	Accuracy: 99.05%
Seal et al. [21]	Deep convolution neural network	EEG	Accuracy: 99.37% AUC: 0.99
Kaur et al. [22]	Variational mode decomposition, Empirical Mode decomposition,	EEG	Accuracy: 99.97%
Mitra et al. [23]	Mel-frequency cepstral coefficients, damped oscillator cepstral coefficients	Speech	Root mean squared error: 8.789
Afshan et al. [24]	Mel Frequency Cepstral Coefficients	Voice features	Accuracy: 94.79% F1-Score: 95.08% Recall: 97.66% Precision: 92.63%
Williamson et al. [25]	Gaussian mixture model, extreme learning machine	Vocal, Facial Biomarkers	Root-mean-squared-error: 8.12 Mean-absolute-error: 6.31
Ooi et al. [26]	Gaussian mixture model	Speech	Accuracy: 73.00% Sensitivity: 79.00% Specificity: 67.00%
Sturim et al. [27]	Gaussian-mixture models	Speech	Confidence: 95.00%
Taguchi et al. [28]	Mel Frequency Cepstral Coefficients	Vocal acoustic features	Sensitivity:77.80% Specificity: 86.10%
Cohn et al. [29]	Active appearance modeling	Facial actions, vocal prosody	Accuracy: 79.00% Likelihood ratio X^2:31.45
Mitra and Shriberg [30]	Neural networks	Speech	Root mean squared error: 7.37 Mean absolute error: 5.87
Williamson et al. [31]	Gaussian mixture model	Vocal features	Root-mean-squared-error: 7.42 Mean-absolute-error: 5.75
Low et al. [32]	Teager energy operator	Speech	Accuracy Female:79.00% Male:87.00%
Seneviratne and Espy-Wilson [33]	Dilated convolutional neural networks	Speech	Accuracy: 91.84% AUC-ROC: 85.46%
Zhang et al. [34]	Mel Frequency Cepstral Coefficients	Voice biomarkers	AUC-ROC: 82.10% Mean-absolute-error: 4.70
Dibeklioglu et al. [35]	Per-frame coding, per-video Fisher-vector based coding	Facial and head movement, vocal prosody	Accuracy Facial movement dynamics: 72.59% Head movement dynamics: 65.25% Voice: 44.44% Head+Face+Voice: 78.67

3 The presented melamine pattern

Molecule structure are accepted as DNA of the materials and can be used to identify the materials. Moreover, these structures are expressed using graphs. Therefore, machine learning model (especially deep learning models) are used in molecule shapes to reach high performance using a nature-inspired model [48]. However, there is no molecule shape for hand-designed feature generator. Hence, a molecule shape based feature generator called Melamine (the molecular structure of melamine is widely known and it is named ChemSpider) pattern is presented.

Hand-crafted feature generation methodology uses various techniques to generate the most relevant features from the

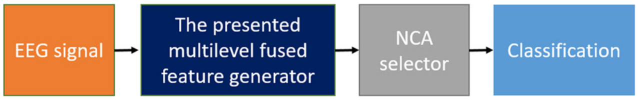

Fig. 1 Proposed melamine pattern and NCA based depression detection model. The collected EEG signals are loaded to system. Multilevel DWT, melamine pattern and statistical features are used to extract both low-level and high-level features. Then, NCA chooses the top 256 features which are used to classify with kNN or SVM classifiers

signals. Histogram based feature generation and local binary pattern (LBP) are the two widely used feature generators [44, 49]. However, LBP has the following limitations [50–52].

- It uses a linear pattern and by employing this pattern, few important features cannot be detected.
- It employs signum function as kernel function which only compares the local values. This may result in losing few important features.

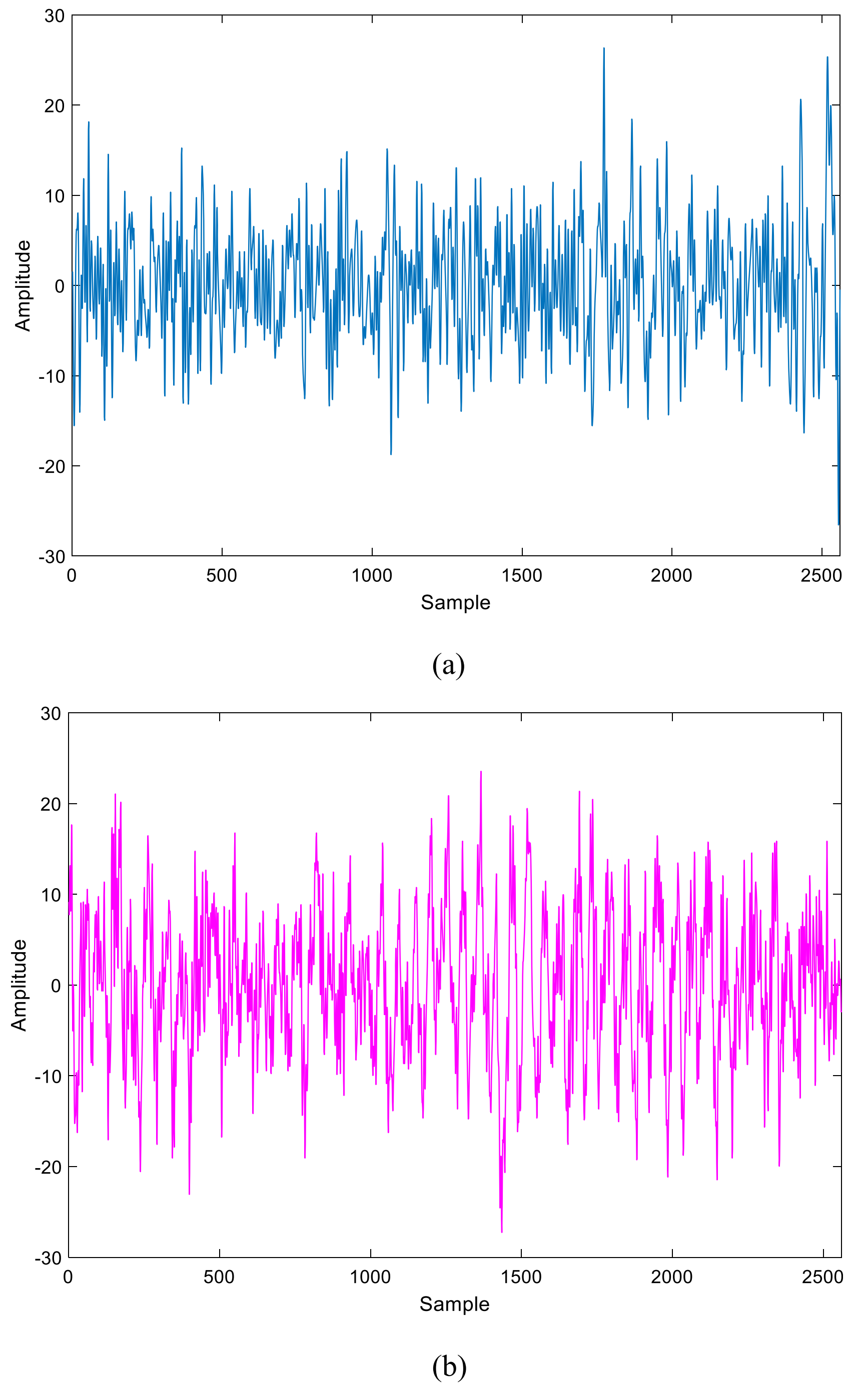

Fig. 2 Sample EEG signal: (**a**) normal, and (**b**) depression

In order to overcome these drawbacks, we proposed two solutions. They are listed below:

- We are proposing to use non-linear patterns for feature generation. In this work, molecular structure of melamine is used as pattern. The shape of molecular structure is called ChemSpider which has the ability to generate unique features.
- Signum function is a good solution for feature generation but it is not able to solve some problems. Hence, we have used *three* nonlinear kernels.

Steps used to obtain the features from melamine pattern is given below.

- **Step 1:** Divide EEG signal into 25 sized overlapping windows.

$$window(j) = ES(i+j-1), i \in \{1,2,\ldots,Ln-24\}, j \in \{1,2,\ldots,25\} \quad (1)$$

Equation (1) defines overlapping block division. Where *window* is 25 sized overlapping block, *Ln* is length of EEG signal (*ES*).

- **Step 2:** Apply vector to matrix transformation and obtain 5 × 5 sized matrix.

$$m(k,l) = window(j), k \in \{1,2,\ldots,5\}, l \in \{1,2,\ldots,5\} \quad (2)$$

where *m* is obtained 5 × 5 sized matrix.

- **Step 3:** Create a pattern using ChemSpider. The typical sketch of ChemSpider is shown in Fig. 3.

Inspired by ChemSpider, a new pattern is presented and graphical sketch of our presented pattern is shown in Fig. 4.

- **Step 4:** Use three kernels for calculating binary features. The used kernels are defined in Eqs. 3–5.

$$k^1(param^1, param^2) = \begin{cases} 0, param^1 - param^2 < 0 \\ 1, param^1 - param^2 \geq 0 \end{cases} \quad (3)$$

$$k^2(param^1, param^2) = \begin{cases} 0, param^1 - param^2 > -param^1 \\ 1, param^1 - param^2 \leq -param^1 \end{cases} \quad (4)$$

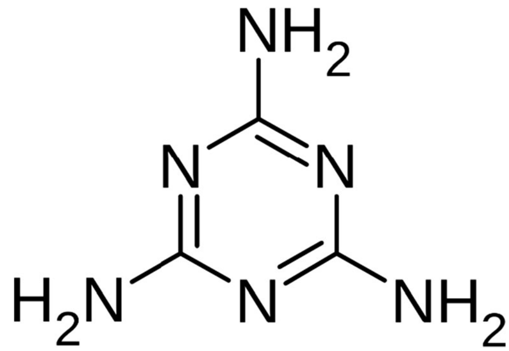

Fig. 3 Molecular structure of melamine (ChemSpider)

$$k^3(param^1, param^2) = \begin{cases} 0, param^1 - param^2 < param^2 \\ 1, param^1 - param^2 \geq param^2 \end{cases} \quad (5)$$

The used kernels are defines as $k^1(.,.)$, $k^2(.,.)$ and $k^3(.,.)$. As seen from Eq. (3)–(5), $k^1(.,.)$ is signum function. $k^2(.,.)$ and $k^3(.,.)$ kernels extract lower and upper signals. Each kernel generates *nine* bits deploying the presented melamine pattern.

- **Step 5:** Extract three binary feature vectors using melamine pattern and the used three kernels. The mathematical explanation of the bit generation process using the defined three kernel and the presented melamine pattern is given in eqs. (6).

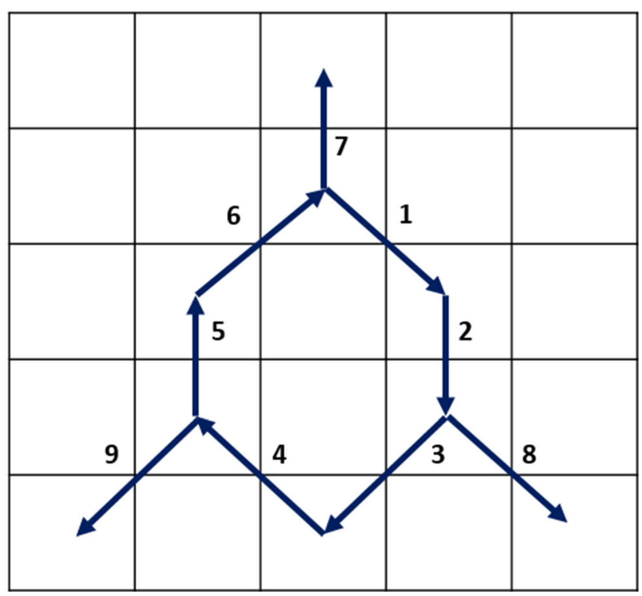

Fig. 4 Melamine pattern created using ChemSpider. The digits represent enumeration of relations. Kernels take two parameters: initial point of each arrow indicates the first parameter and final point of each arrow shows the second parameter of used kernel

$$\begin{bmatrix} bit^h(1) \\ bit^h(2) \\ bit^h(3) \\ bit^h(4) \\ bit^h(5) \\ bit^h(6) \\ bit^h(7) \\ bit^h(8) \\ bit^h(9) \end{bmatrix} = k^h \left(\begin{bmatrix} m(2,3), m(3,4) \\ m(3,4), m(4,4) \\ m(4,4), m(5,3) \\ m(5,3), m(4,2) \\ m(4,2), m(3,2) \\ m(3,2), m(2,3) \\ m(2,3), m(1,3) \\ m(4,4), m(5,5) \\ m(4,2), \mathrm{m}(5,1) \end{bmatrix} \right), h \in \{1,2,3\} \quad (6)$$

where bit^1, $bit^2$and bit^3 are generated bits using $k^1(.,.)$, $k^2(.,.)$ and $k^3(.,.)$ respectively. Equation (6) denotes the bit generation processes. Each kernel generates nine bits using melamine pattern. For instance, the first bit of the upper signal ($bit^{upper}(1)$) is generated using $k^3(m(2,3), m(3,4))$.

- **Step 6:** Binary to decimal conversion is applied on the generated bits.

$$SS(i) = \sum_{j=1}^{9} bit^1(j) * 2^{j-1}, \quad i \in \{1, 2, \ldots, Ln-24\} \quad (7)$$

$$SL(i) = \sum_{j=1}^{9} bit^1(j) * 2^{j-1} \quad (8)$$

$$SU(i) = \sum_{j=1}^{9} bit^1(j) * 2^{j-1} \quad (9)$$

where SS, SL and SU represent signum signal, lower signal and upper signal consecutively.

- **Step 7:** Extract histograms of these signals. Each signal coded with 9-bits. Therefore, the size of generated each histogram is $2^9 = 512$.
- **Step 8:** Merge the generated histograms and calculate feature vectors of length = 1536.

The given eight steps define our melamine pattern feature generator.

4 Proposed method

This work proposes a hand-crafted feature for automated detection of MDD using EEG signals. Therefore, a signal processing model is presented and this model has *three* main steps. In the first step, the recommended melamine pattern and DWT are used to extract the features. 25 statistical features are extracted from the presented model. The proposed feature generator extracts two types features. Melamine pattern generates textural features and the 25 statistical moments are extracted from statistical features. The six-leveled DWT is performed using Daubechies four (db4) mother wavelet function [53, 54]. The generated features are fed to NCA [55] and the most discriminative 256 features are selected. The selected features are classified using SVM and kNN classifiers developed using 80:20 hold-out cross validation strategy. The snapshot of the proposed automated major depression detection model is shown in Fig. 5.

Moreover, pseudocode of the introduced Melamine pattern and NCA based MDD classification model is shown in Algorithm 1.

Algorithm 1. Pseudocode of the recommended melamine pattern and NCA based MDD model.

Input parameters: *EEG based MDD dataset (D) with a size of 20 x 7339 x 2560 (number of channels x number of observations x length of EEG record)*
Output: *Results*

```
for c=1 to 20 do // Read channel
    for k=1 to 7339 do // Read signal.
        signal = D_{c,k};
        Apply six levels DWT to signal.
            Extract features deploying the presented melamine pattern and the used statistics.
        Merge the extracted features.
    end for k
    Deploy NCA and select the most valuable features.
    Classify the chosen features and obtain results for the used channel.
end for c
```

The steps of the presented melamine pattern and NCA-based MDD model is given below.

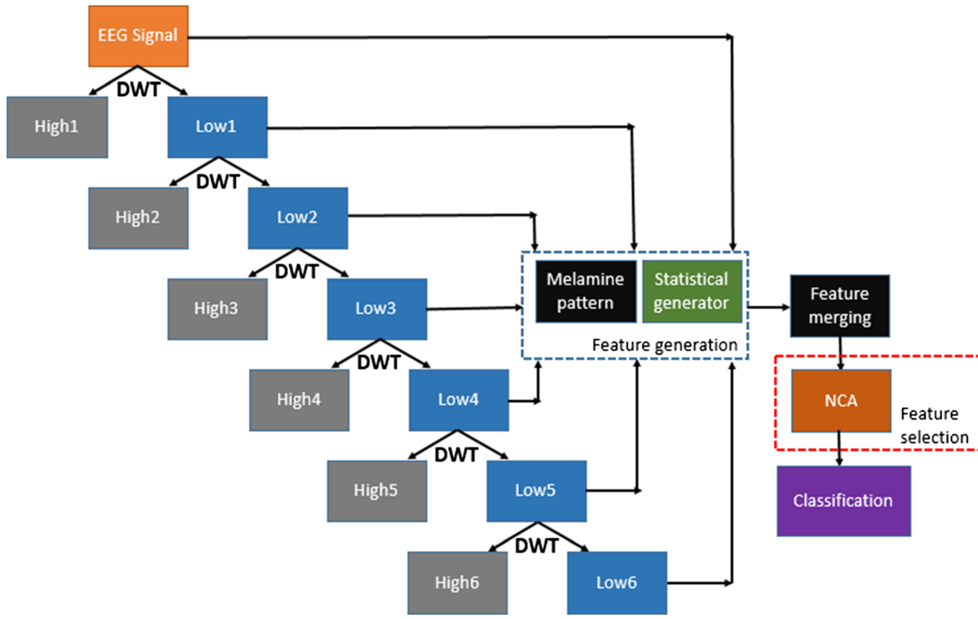

Fig. 5 Snapshot of the proposed automated major depression detection model

4.1 Feature generation

In each level of DWT, statistical feature generator and the presented melamine pattern are utilized together. The primary objective of this step is to extract features of both low and high levels. The DWT generates both levels (low and high frequencies), statistical feature generator extracts statistical features and proposed melamine pattern extracts textural features from the DWT coefficients. The steps involved to perform the features extraction are given below:

- **Step 0:** Load EEG signal.
- **Step 1:** Apply six-level DWT on the EEG signal.

$$NL = \left\lfloor \log_2\left(\frac{Ln}{25}\right) \right\rfloor \tag{10}$$

Equation (10) defines the number of DWT levels. Melamine pattern uses 25 sized overlapping signal and length of the EEG signal used is 2560. Therefore, $NL = \left\lfloor \log_2\left(\frac{2560}{25}\right) \right\rfloor = 6$ is calculated as number of levels.

$$\left[low^1, high^1\right] = DWT(ES, db4) \tag{11}$$

$$\left[low^k, high^k\right] = DWT\left(low^{k-1}, db4\right), k\in\{2, 3, \ldots, 6\} \tag{12}$$

where low^k and $high^k$ are k$^{\text{th}}$ low-pass and high-pass filters coefficients, $DWT(.,.)$ function represent one dimensional DWT.

- **Step 2:** Generates textural features from EEG signal and features are generated using melamine patterns using low-pass coefficients of DWT.

$$tf^1 = MP(ES) \tag{13}$$

$$tf^{t+1} = MP(low^t), t\in\{1, 2, \ldots, 6\} \tag{14}$$

where tf^t is t$^{\text{th}}$ level textural feature vector and $MP(.)$ defines the presented melamine pattern and it generates 1536 features.

- **Step 3:** Extract 25 statistical features from DWT coefficients.

$$sf^1 = St(ES) \tag{15}$$

$$sf^{t+1} = St(low^t), t\in\{1, 2, \ldots, 6\} \tag{16}$$

where sf^t is t$^{\text{th}}$ level statistical features, $St(.)$ is used statistical feature generation function and uses 25 statistical moments. Therefore, the length of each sf is 25. The used statistical moments are given in Table 2 [56].

These moments are consisted of the $St(.)$ Feature generation function.

- **Step 4:** Merge the generated statistical and textural features.

$$X(1561*k + j) = conc\left(tf^{k+1}, sf^{k+1}\right), k\in\{0, 1, \ldots, 6\}, j\in\{1, 2, \ldots, 1561\} \tag{17}$$

Table 2 Details of 25 statistical moments extracted

Number	Moment	Number	Moment
1	Mean	14	Shannon entropy
2	Standard deviation	15	Sure entropy
3	Variance	16	Log energy entropy
4	Median	17	Absolute Shannon entropy
5	Maximum	18	Absolute sure entropy
6	Minimum	19	Absolute log energy entropy
7	Range (Max-min)	20	Kurtosis
8	Root mean square	21	Absolute kurtosis
9	Energy	22	Skewness
10	Absolute mean	23	Absolute skewness
11	Absolute standard deviation	24	Mean square deviation
12	Absolute variance	25	Mean absolute deviation
13	Absolute median		

where X defines the generated and merged statistical and textural features with a length of $1561 * 7 = 10927$.

4.2 Feature selection

Feature selection is one of the most critical steps of classification. The presented melamine pattern and NCA based model used a multileveled feature generation model. Therefore, 10,927 features (it is huge feature set) are generated using the presented multileveled (DWT- based) hybrid feature generation model. The size of this feature vector must be decreased. The general objectives of the feature selection methods are [57, 58]: (i) selection the most valuable features to increase the performance, (ii) decreasing number of features decrease the execution times of the classifiers. To meet these requirements, NCA is chosen as feature selector. The NCA is one of the simplest feature selector and it is a feature selection variance of the nearest neighborhood model [59, 60]. It generates non-negative (positive) feature weights which helps to select the most discriminative features. The steps involved are given below:

- **Step 5:** Apply NCA to generate and merge the features (X).
- **Step 6:** Select the most discriminative 256 features by using the generated weights.

4.3 Classification

The selected 256 features are fed to Quadratic SVM and Weighted kNN classifiers for automated classification. To select the most appropriate classifier, MATLAB classification learner toolbox (MCLT) which has 25 classifiers is used in this work. We have obtained highest classification accuracy using quadratic SVM and weighted kNN classifiers. The attributes of the classifiers used are given as below.

- ***Weighted kNN:*** In this work, k value is selected as 10, distance metric is spearman and distance weight is selected as squared inverse [61].
- ***Quadratic SVM:*** In this work, we have chosen second-degree polynomial kernel, box constraint level = 1 and one-vs-one as multiple classification model [45, 46] to obtain the maximum performance. We have developed the model using hold-out validation strategy with 80% is used for training and 20% for testing.
- **Step 7:** Classify the selected features using Quadratic SVM or weighted kNN classifiers.

5 Experimental results

The presented melamine pattern and NCA- based model is implemented by using MATLAB (2020a) programming environment and it is programmed functionally. The used functions are named as melamine pattern, statistical feature generator, NCA and main functions. The main function and feature generation and selection functions generate features. The selected features are fed to MATLAB classification learner application containing 25 classifiers. We have used quadratic SVM and Weighted kNN classifiers in this work. We have developed the classification model for every channel using 80:20 hold-out validation strategy. To evaluate performance of this model for each channel, sensitivity, specificity, balanced accuracy, overall accuracy, geometric mean, F1-score and precision metrics were used. True positives (TP), false positives (FP), true negative (TN) and false negatives (FN) parameters are computed. Mathematical notations of the used measurements were given below (Eqs. (18)–(24)) [62, 63].

$$sen = \frac{TP}{TP + FN} \quad (18)$$

$$spe = \frac{TN}{TN + FP} \quad (19)$$

$$gm = \sqrt{sen*spe} \quad (20)$$

$$bacc = \frac{sen + spe}{2} \quad (21)$$

$$acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (22)$$

$$pre = \frac{TP}{TP + FP} \quad (23)$$

$$F1 = 2\frac{sen*pre}{sen + pre} \quad (24)$$

where *sen*, *spe*, *gm*, *bacc*, *acc*, *pre* and $F1$ defines sensitivity, specificity, geometric mean, balanced accuracy, accuracy, precision and F1 score respectively. We have developed the model using 80:20 hold-out validation strategy and classifiers are executed 100 times. The minimum, average, maximum and standard deviation values of the results are listed in Tables 3-6 using two classifiers with various channels. There are 20 channels in the used EEG dataset and results of ten channels are given in each table. Furthermore, best results are highlighted in bold font in these tables.

The obtained results for different channels using quadratic SVM is given in Tables 3-4.

In this Table, Min, Av., Max and Std define the minimum, maximum, average and standard deviation values respectively. In Tables 3, 99.05% classification accuracy is achieved using A2A1 channel. The general classification accuracy of this channel is 98.23% ± 0.32%. The worst channel is F7, as it reached 95.57% maximum accuracy using SVM classifier.

Table 3 Summary of performance measures (%) obtained using our proposed system with quadratic SVM classifier for 1st -10th EEG channels

Metric	St.	A2A1	C3	C4	Cz	F3	F4	F7	F8	Fp1	Fp2
Acc(%)	Min	97.14	93.52	93.87	93.66	95.02	93.66	92.91	92.91	93.25	93.93
	Av.	98.23	95.28	95.09	95.12	96.04	95.10	94.43	94.51	94.60	95.21
	Max	**99.05**	96.73	96.39	96.52	97.68	96.52	95.57	96.11	95.77	96.39
	Std	0.32	0.50	0.52	0.52	0.48	0.58	0.51	0.57	0.53	0.49
Sen(%)	Min	95.50	92.16	91.58	91.0	92.89	90.71	89.40	89.70	89.70	90.57
	Av.	97.53	94.56	93.97	93.78	95.56	93.18	92.34	92.14	92.10	93.47
	Max	**99.27**	96.23	95.94	96.08	97.24	95.36	94.10	95.21	94.05	95.36
	Std	0.57	0.83	0.88	0.88	0.89	0.94	0.98	0.98	0.95	1.00
Spe(%)	Min	97.56	93.70	94.09	94.34	94.22	94.34	94.60	95.37	95.24	94.86
	Av.	98.85	95.91	96.09	96.31	96.47	96.79	96.29	96.61	96.81	96.76
	Max	**99.87**	97.96	97.94	98.71	98.07	98.59	97.69	98.33	98.33	98.33
	Std	0.38	0.67	0.75	0.71	0.73	0.62	0.68	0.69	0.65	0.63
gm(%)	Min	97.07	93.51	93.78	93.52	95.07	93.50	92.65	92.68	93.0	93.80
	Av.	98.19	95.23	95.02	95.04	96.01	94.97	94.29	94.35	94.42	95.10
	Max	**99.04**	96.66	96.32	96.42	97.66	96.41	95.46	96.06	95.66	96.31
	Std	0.33	0.51	0.53	0.53	0.49	0.55	0.53	0.58	0.55	0.52
bacc(%)	Min	97.08	93.51	93.79	93.54	95.08	93.53	92.71	92.73	93.06	93.82
	Av.	98.19	95.23	95.03	95.05	96.01	94.99	94.31	94.38	94.46	95.11
	Max	**99.04**	96.67	96.33	96.43	97.66	96.42	95.47	96.06	95.68	96.32
	Std	0.33	0.51	0.53	0.53	0.49	0.54	0.52	0.58	0.54	0.51
Pre(%)	Min	97.25	92.92	93.36	93.59	93.63	93.60	93.76	94.63	94.49	94.22
	Av.	98.69	95.35	95.52	95.75	96.0	96.26	95.67	96.02	96.24	96.24
	Max	**99.85**	97.34	95.78	98.47	97.81	98.32	97.19	98.02	97.97	98.01
	Std	0.43	0.72	0.82	0.78	0.78	0.69	0.75	0.78	0.73	0.69
F1(%)	Min	96.93	93.12	93.41	93.14	94.77	93.12	92.22	92.24	92.60	93.44
	Av.	98.11	94.95	94.73	94.75	95.78	95.10	93.97	94.04	94.12	94.83
	Max	**98.98**	96.49	96.12	96.25	97.53	96.52	95.22	95.84	95.43	96.12
	Std	0.35	0.54	0.56	0.57	0.48	0.53	0.51	0.63	0.59	0.55

Table 4 Summary of performance measures (%) obtained using our proposed system with quadratic SVM classifier for 11th -20th EEG channels

Metric	St.	Fz	O1	O2	P3	P4	Pz	T3	T4	T5	T6
Acc(%)	Min	94.27	94.27	93.93	94.07	94.61	93.93	92.98	95.84	93.46	94.14
	Av.	95.45	95.47	95.12	95.43	95.79	95.20	94.65	96.64	95.46	95.33
	Max	96.86	96.32	96.39	96.80	97.0	96.39	96.05	**97.61**	96.52	96.39
	Std	0.50	0.46	0.53	0.51	0.45	0.54	0.58	0.42	0.55	0.47
Sen(%)	Min	91.73	90.42	89.99	91.73	92.02	90.71	88.68	93.47	90.86	91.0
	Av.	93.84	92.75	92.26	93.93	94.17	93.38	91.74	95.40	93.50	93.35
	Max	95.94	94.34	94.92	95.94	96.23	96.08	94.05	**97.24**	96.37	95.07
	Std	0.87	0.84	1.03	0.90	0.84	0.71	1.08	0.82	1.04	0.85
Spe(%)	Min	95.24	96.40	96.02	95.24	95.76	95.12	96.02	96.27	95.24	95.37
	Av.	96.89	97.88	97.66	96.77	97.24	96.81	97.23	97.74	97.19	97.08
	Max	95.94	99.10	98.59	98.07	98.97	98.33	98.33	**98.84**	98.71	98.71
	Std	0.87	0.55	0.51	0.67	0.62	0.64	0.51	0.52	0.70	0.65
gm(%)	Min	94.10	93.98	93.67	93.98	94.49	93.71	92.66	95.78	93.32	93.92
	Av.	95.35	95.28	94.92	95.34	95.69	95.08	94.44	96.56	95.32	95.19
	Max	96.79	96.19	96.29	96.71	96.90	96.31	95.87	**97.58**	96.51	96.28
	Std	0.52	0.48	0.56	0.52	0.46	0.56	0.62	0.44	0.57	0.49
bacc(%)	Min	94.13	94.05	93.73	93.99	94.51	93.76	92.74	95.78	93.34	93.97
	Av.	95.36	95.31	94.96	95.35	95.70	95.10	94.48	96.57	95.34	95.21
	Max	96.79	96.21	96.30	96.72	96.91	96.31	95.90	**97.58**	96.51	96.29
	Std	0.51	0.47	0.55	0.52	0.46	0.55	0.61	0.44	0.56	0.48
Pre(%)	Min	94.58	95.82	95.35	94.63	95.17	94.48	95.30	95.82	94.45	94.72
	Av.	96.39	97.49	97.22	96.27	96.80	96.29	96.71	97.40	96.72	96.60
	Max	97.63	98.91	98.32	97.74	98.78	97.99	98.02	**98.65**	98.48	98.46
	Std	0.64	0.64	0.59	0.74	0.69	0.71	0.59	0.58	0.79	0.72
F1(%)	Min	93.77	93.68	93.33	93.62	94.18	93.36	92.24	95.54	92.92	93.59
	Av.	95.09	95.06	94.67	95.08	95.79	94.81	94.15	96.38	95.08	94.94
	Max	96.63	96.01	96.11	96.55	96.76	96.11	95.69	**97.45**	96.30	96.10
	Std	0.55	0.51	0.60	0.55	0.49	0.59	0.66	0.46	0.60	0.52

It can be noted from Table 4 that, A2A1 channel reached the highest accuracy among all channels. The best accurate channel is T4 and it reached 97.61% maximum accuracy and general accuracy of 96.64% ± 0.42%. There is no channel with accuracy lower than F7 channel in Table 4. The channel-wise results obtained using weighted kNN are listed in Tables 5-6.

It can be noted from Table 5 that, 99.11% classification accuracy is achieved using A2A1 channel. The general classification accuracy of this channel is 98.29% ± 0.34%. The channel F7 reached the worst results with maximum accuracy of 95.53% using Weighted kNN classifier with general results = 93.39% ± 0.34%.

It can be seen from these results that, highest classification accuracy of 98.57% and 99.11% are obtained for channels A2A1 and T4 channels, respectively using weighted kNN classifier. Our same model obtained the accuracy of 99.05% and 97.61% using A2A1 and T4 channels, respectively with quadratic SVM classifier. In both classifiers, the best accurate channel is found to be A2A1 using both classifiers and F7 channel yielded the worst results for both classifiers. Tables 3-6 denotes the results of our proposed melamine pattern and NCA based model reached accuracies of >95% for all channels. Also, our MDD model reached specificity of 99.87% and recall rate of 99.85% for A2A1 channel using weighted kNN classifier.

Performance measures (sensitivity, specificity, geometric mean, precision and F1-score) obtained for various channels using SVM, and (b) weighted kNN classifiers are shown in Fig. 6.

6 Discussions

The primary motivation of this work is to develop a novel feature generation model using molecular structure graph called melamine pattern. Our proposed model

Table 5 Summary of performance measures (%) obtained using our proposed system with Weighted kNN classifier for 1st -10th EEG channels

Metric	St.	A2A1	C3	C4	Cz	F3	F4	F7	F8	Fp1	Fp2
Acc(%)	Min	97.14	94.75	94.0	93.59	94.82	93.73	92.57	92.64	93.39	95.02
	Av.	98.29	96.05	95.46	95.21	96.14	95.31	93.39	94.36	94.59	96.13
	Max	**99.11**	97.48	96.80	96.73	97.34	96.86	95.23	95.64	96.11	97.34
	Std	0.34	0.50	0.51	0.54	0.50	0.56	0.53	0.57	0.56	0.48
Sen(%)	Min	95.07	90.57	89.55	90.28	91.58	89.40	88.24	87.08	86.79	91.73
	Av.	96.92	92.91	92.65	92.74	94.07	91.95	90.52	90.46	89.96	93.86
	Max	**98.40**	95.94	94.92	95.07	96.23	94.78	92.89	92.60	92.89	96.23
	Std	0.64	0.97	0.95	0.96	0.91	1.01	0.92	1.11	1.10	0.90
Spe(%)	Min	98.33	97.56	96.40	95.89	96.66	96.66	95.63	96.66	97.81	96.92
	Av.	99.32	98.82	97.96	97.39	97.97	98.29	97.03	97.81	98.69	98.14
	Max	**99.87**	99.61	98.97	98.97	99.10	99.49	98.33	98.84	99.74	99.23
	Std	0.30	0.35	0.47	0.56	0.50	0.47	0.51	0.48	0.39	0.46
gm(%)	Min	97.02	94.46	93.71	93.39	94.63	93.41	92.31	92.22	92.80	94.78
	Av.	98.11	95.82	95.26	95.04	96.0	95.07	93.73	94.06	94.22	95.97
	Max	**99.06**	97.38	96.67	96.59	97.23	96.72	95.01	95.42	95.84	97.22
	Std	0.36	0.54	0.54	0.56	0.52	0.59	0.56	0.61	0.61	0.51
bacc(%)	Min	97.04	94.53	93.78	93.43	94.67	93.49	92.37	92.35	93.0	94.84
	Av.	98.12	95.87	95.30	95.07	96.02	95.12	93.78	94.14	94.32	96.0
	Max	**99.06**	97.39	96.69	96.61	97.24	96.74	95.03	95.46	95.91	97.24
	Std	0.36	0.52	0.53	0.55	0.52	0.58	0.55	0.59	0.59	0.50
pre(%)	Min	98.10	97.11	95.78	95.30	96.13	96.04	94.75	95.99	97.33	96.43
	Av.	99.21	98.59	97.57	96.93	97.63	97.94	96.44	97.35	98.38	97.82
	Max	**99.85**	99.53	98.76	98.77	98.92	99.37	97.98	98.55	99.68	99.08
	Std	0.34	0.42	0.54	0.64	0.57	0.56	0.60	0.57	0.47	0.53
F1	Min	96.90	94.21	93.38	93.01	94.35	93.06	91.85	91.78	92.50	94.54
	Av.	98.05	95.66	95.04	94.78	95.81	94.85	93.39	93.77	93.98	95.79
	Max	**99.05**	97.28	96.53	96.45	97.13	96.60	94.75	95.22	95.72	97.12
	Std	0.37	0.57	0.58	0.60	0.56	0.63	0.60	0.65	0.65	0.54

consists of melamine pattern, statistical feature generator, DWT, NCA selector and two conventional classifiers. In this work, we have used a huge dataset consisting of 64 subjects with 20 channels of EEG signals for each subject. The used testbed (corpus) has 20 channels with 7339 EEG signals in each channel. We have used seven performance measures to evaluate the performance of our model. The graph of accuracy (%) obtained versus various EEG channels for SVM and kNN classifiers is shown in Fig. 7.

The Fig. 7 demonstrates that the weighted kNN attained higher results than quadratic SVM for more number of channels. However, quadratic SVM reached higher results than Weighted kNN for 5th (F3), 7th (F7), 8th (F8) and 16th (Pz) channels. The highest accuracies are obtained using A2A1 and T4 channels. Confusion matrices of these channels are also shown in Fig. 8.

The summary of state-of-the-art techniques developed for automated detection of depression using EEG signals is shown in Table 7.

It can be seen from Table 7 that the presented model has obtained the high classification performance for depression detection. The presented model is a hand-modeled feature extraction based machine learning model and it selects the most informative features using NCA. Table 7 denotes that the deep learning based models have reached highest performance using big MDD datasets and hand-crafted models achieved high performance on the smaller MDD corpora. The best performing model among state-of-the-art classification methods is Sandheep et al.'s [75] method and they used one-dimensional CNN to classify 6000 EEG records using ten-fold cross-validation and their model achieved 99.31% classification accuracy. However, they reached this result using all channels of the dataset. Mumtaz and Qayyum [74] presented a EEG segmentation, one-

Table 6 Summary of performance measures (%) obtained using our proposed system with Weighted kNN classifier for 11th -20th EEG channels

Metric	St.	Fz	O1	O2	P3	P4	Pz	T3	T4	T5	T6
Acc(%)	Min	94.34	94.41	94.34	94.96	94.48	93.32	93.12	96.52	95.30	94.27
	Av.	95.72	95.68	95.70	96.18	96.0	94.90	94.99	97.46	96.18	95.62
	Max	97.07	97.34	97.0	97.41	97.55	96.32	96.18	**98.57**	97.96	96.80
	Std	0.53	0.46	0.47	0.50	0.50	0.54	0.56	0.39	0.48	0.55
Sen(%)	Min	89.70	88.68	88.82	90.71	90.71	87.81	87.66	94.63	91.29	89.11
	Av.	92.64	91.70	91.57	92.97	92.91	90.68	91.23	96.62	93.64	91.63
	Max	95.21	95.21	94.34	95.50	95.50	93.32	93.76	**98.11**	95.94	94.19
	Std	1.04	0.94	0.95	1.01	0.95	1.09	1.09	0.63	0.54	1.14
Spe(%)	Min	96.92	97.94	98.46	97.81	97.17	97.43	96.27	97.04	97.30	98.33
	Av.	98.44	99.20	99.36	99.01	98.75	98.64	98.32	98.20	98.43	99.15
	Max	**99.36**	99.87	100.0	99.61	99.74	99.74	99.36	**99.36**	99.74	99.87
	Std	0.45	0.31	0.26	0.34	0.43	0.43	0.51	0.47	0.47	0.31
gm(%)	Min	94.10	93.93	93.88	94.67	94.21	92.86	92.66	96.45	94.99	93.85
	Av.	95.49	95.37	95.38	95.94	95.78	94.57	94.71	97.41	96.0	95.31
	Max	96.95	97.20	96.73	97.27	97.39	96.11	95.95	**98.51**	97.82	96.62
	Std	0.60	0.50	0.52	0.54	0.53	0.59	0.60	0.40	0.51	0.60
bacc(%)	Min	94.15	94.08	94.03	94.74	94.27	93.0	92.80	96.46	95.07	93.98
	Av.	95.54	95.45	95.47	95.99	95.83	94.66	94.78	97.41	96.04	95.39
	Max	96.96	97.22	96.85	97.29	97.41	96.15	95.98	**98.52**	97.84	96.65
	Std	0.55	0.48	0.50	0.52	0.52	0.57	0.58	0.40	0.51	0.58
pre(%)	Min	96.34	97.56	98.10	97.44	96.69	96.88	95.58	96.66	96.89	98.0
	Av.	98.14	99.02	99.22	98.82	98.50	98.33	97.97	97.95	98.15	98.96
	Max	99.23	99.84	100.0	99.54	99.69	99.68	99.22	**99.26**	99.70	99.84
	Std	0.53	0.34	0.31	0.40	0.50	0.52	0.60	0.53	0.54	0.37
F1(%)	Min	93.79	93.71	93.65	94.44	93.92	92.51	92.28	96.26	94.80	93.60
	Av.	95.3	95.22	95.24	95.80	95.62	94.35	94.47	97.28	95.84	95.62
	Max	96.83	97.11	96.73	97.19	97.33	95.97	95.80	**98.46**	97.78	96.80
	Std	0.60	0.53	0.54	0.56	0.56	0.62	0.64	0.42	0.54	0.55

dimensional CNN and two-layered LSTM based MDD classification model. They used 19 channels (all channels) of the used dataset and their model (1DCNN-LSTM) attained 98.32% classification rate, but the time burden of this model is very high. Sharma et al. [73] presented a hand-designed features based MDD classification model and reached 99.58% classification accuracy using small dataset. Moreover, they did not give channel-wise results. Our presented model achieved 99.11% and 99.14% accuracies using 80:20 split ratio and ten-fold cross-validation strategy, respectively. Our presented results are single channel results and the performance of the model is given for every channel. Furthermore, the presented model is developed using big MDD dataset and majority voting or channel concatenation is not employed to reach higher performance.

The advantages of this work are as follows:

– A new public dataset is used to develop the machine learning model and obtain the highest classification performance.
– A new feature generation model is presented for textural features using melamine pattern.
– The effective and highly accurate model is developed.
– Model attained >95% accuracies for all channels of EEG signals. These results imply that the developed model is robust and do not need to use all channels.
– This model is lightweight and need not set millions of parameters like deep networks to attain highest classification accuracy.
– This model used an effective feature generation method with a low computation complexity. This feature generation model used multiple kernel based melamine pattern. The time complexity of the melamine pattern is calculated as $O(n)$. Moreover, the one dimensional DWT is used to decompose signals and it halves length of the signal in each level. Therefore, the time complexity of the presented multileveled feature generation model is calculated as $O(nlogn)$.
– Developed a simple model and can be implemented easily as it does not involve lot of calculations.

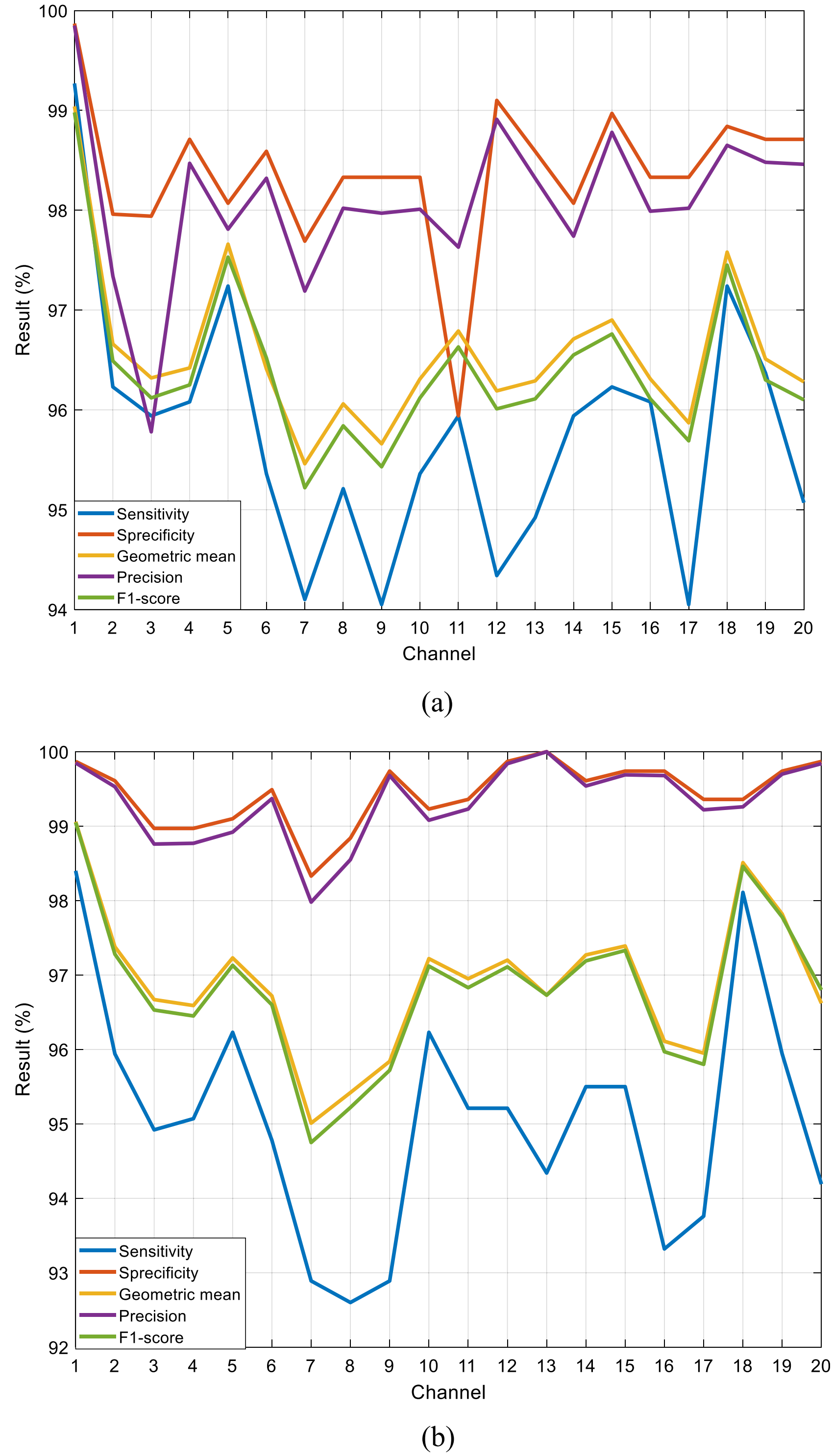

Fig. 6 Performance measures (sensitivity, specificity, geometric mean, precision and F1-score) obtained for various channels using classifier: (**a**) SVM, and (**b**) weighted kNN

In this work, melamine pattern and NCA-based method is employed to detect MDD automatically using EEG signals. There are many molecules in nature and they can be defined as a graph. More molecular structure based feature generation models can be presented to improve the classification performance. The developed automated MDD detection model can be uploaded to the cloud. The EEG signals to be tested is sent to our model placed in the cloud to find the class of the EEG signal. The result of the model will be sent to the clinicians and after confirming from the psychiatrists will be sent to the patients. This will expedite the diagnosis process and also help to provide immediate treatment to the patients.

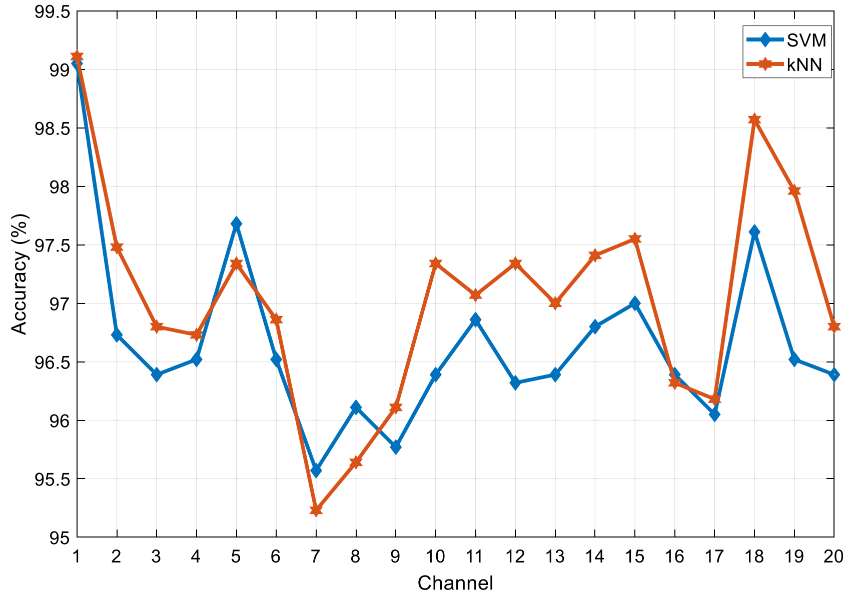

Fig. 7 Graph of accuracy (%) obtained versus various EEG channels for SVM and kNN classifiers

7 Conclusions

Automated depression detection using EEG signals is a complex and challenging problem in machine learning. Many deep learning models have been used to attain high performance. In this work a hand-crafted features-based depression detection model is proposed using novel melamine pattern with EEG signals. Our presented model attained >95% accuracies for all 20 channels of EEG signals. This clearly indicates the robustness of the developed system. We have obtained highest classification accuracy of 99.11% and 99.05% using Weighted kNN and Quadratic SVM respectively using A2A1 channel. In future, the developed model can be used to detect early (mild) stage of depression using EEG signals using bigger database.

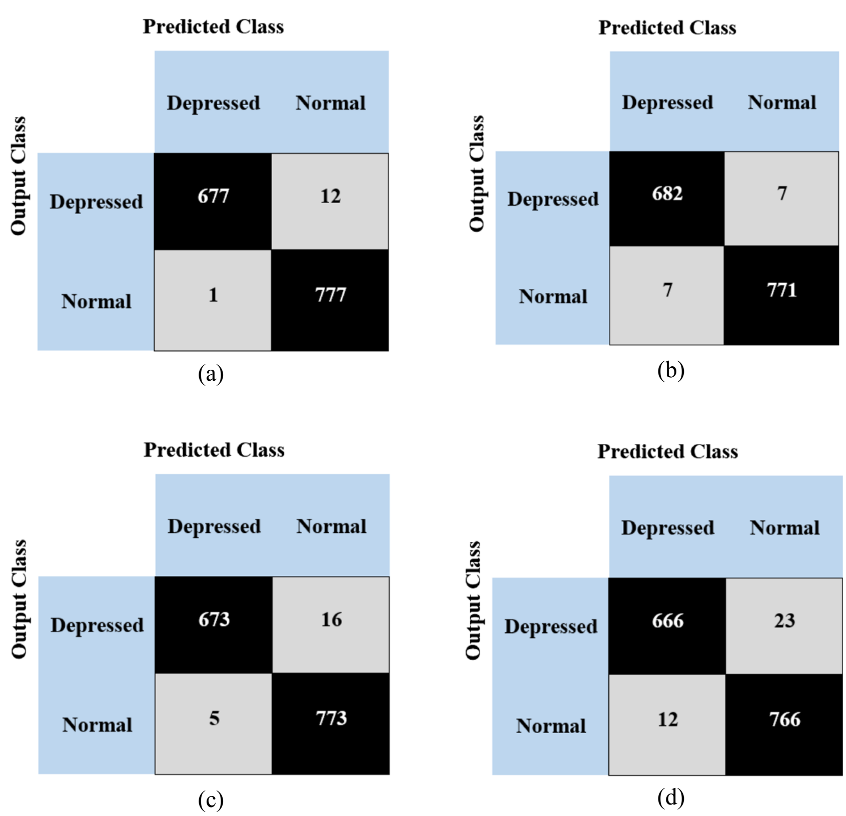

Fig. 8 Confusion matrices of the best performing channels using kNN and SVM classifiers. These matrices denote number of true predicted and false predicted values. **a** Confusion matrix obtained using A2A1 channel with kNN classifier. **b** Confusion matrix obtained using A2A1 channel with SVM classifier. **c** Confusion matrix obtained using T4 channel with kNN classifier. **d** Confusion matrix obtained using T4 channel with SVM classifier

Table 7 Summary of state-of-the-art techniques developed for automated detection of depression using EEG signals

Study	Year	Method	Data	Accuracy	Number of observations	Duration of the used EEG records
Mantri et al. [64]	2015	Fast Fourier Transform	12 healthy, 13 major depression	84.00	–	5 min
Acharya et al. [65]	2015	Recurrence Quantification Analysis, Higher Order Spectra	15 healthy, 15 major depression	98.00	–	5 min
Erguzel et al. [66]	2016	Artificial neural network, Particle swarm optimization	31 bipolar subjects 58 unipolar subjects	89.89	–	2 min
Mumtaz et al. [67]	2017	Wavelet transform	30 healthy, 34 major depression	87.50	53,675	1,2,3 min
Liao et al. [68]	2017	Common spatial pattern	12 healthy, 12 major depression	81.23	–	5 min
Kim et al. [69]	2018	Support vector machine recursive feature elimination	37 healthy, 30 major depression	74.00	–	150 s
Cai et al. [70]	2018	Discrete Wavelet Transformation, Adaptive Predictor Filter	121 healthy, 92 major depression	79.27	–	12 s
Wu et al. [71]	2018	Conformal kernel support vector machine	31 healthy, 24 major depression	83.64	–	6 s
Acharya et al. [72]	2018	Convolutional neural network	15 healthy, 15 major depression	95.49 (Right) 93.96 (Left)	4333 Healthy, 4333 MDD	5 min
Sharma et al. [73]	2018	Three-channel orthogonal wavelet filter bank, bandwidth-duration localization	15 healthy, 15 major depression	99.58	4200 Healthy, 4200 MDD	5 min
Mahato and Paul [4]	2019	Multi layered perceptron neural network	30 healthy, 34 major depression	93.33	–	5 min
Mumtaz and Qayyum [74]	2019	One dimensional convolutional neural network	30 healthy, 33 major depression	98.32	–	5 min
Sandheep et al. [75]	2019	Convolutional neural network	30 healthy, 30 depression	99.31	3000 Healthy, 3000 MDD	–
Li et al. [76]	2019	Convolutional neural network	27 healthy, 24 depression	85.62	–	6 s
Mohammadi et al. [77]	2019	Fuzzy function based on neural network	60 participants (severe=23, minimal=17, mild=10, moderate=10)	87.50	–	4 min
Ay et al. [78]	2019	Convolutional neural network, long-short term memory	15 healthy, 15 major depression	99.12 (Right) 97.66 (Left)	4318 Healthy, 4798 MDD	7.8 s
Duan et al. [79]	2020	Convolutional neural network	16 healthy, 16 major depression	94.13	29,229 Healthy 35,242 MDD	1,2,3 s
Our Method		**Melamine pattern**	**30 healthy,** **34 major depression**	**99.11 with 80:20 split ratio**	**3893 Healthy and 3446 MDD**	**10 s**
Our Method		**Melamine pattern**	**30 healthy,** **34 major depression**	**99.14 with 10-fold cross-validation**	**3893 Healthy and 3446 MDD**	**10 s for each segment**

References

1. Belmaker R, Agam G (2008) Major depressive disorder. N Engl J Med 358:55–68
2. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M et al (2016) Major depressive disorder. Nature reviews Disease primers 2:1–20
3. Lohoff FW (2010) Overview of the genetics of major depressive disorder. Current psychiatry reports 12:539–546
4. Mahato S, Paul S (2019) Detection of major depressive disorder using linear and non-linear features from EEG signals. Microsyst Technol 25:1065–1076
5. Lehman JF. The diagnostic and statistical manual of mental disorders. 2000
6. Yasin S, Hussain SA, Aslan S, Raza I, Muzammel M, Othmani A. Neural Networks based approaches for Major Depressive Disorder and Bipolar Disorder Diagnosis using EEG signals: A review. arXiv preprint arXiv:200913402. 2020
7. Stockings E, Degenhardt L, Lee YY, Mihalopoulos C, Liu A, Hobbs M, Patton G (2015) Symptom screening scales for detecting major depressive disorder in children and adolescents: a systematic review and meta-analysis of reliability, validity and diagnostic utility. J Affect Disord 174:447–463
8. Akar SA, Kara S, Agambayev S, Bilgiç V (2015) Nonlinear analysis of EEGs of patients with major depression during different emotional states. Comput Biol Med 67:49–60
9. Landsness EC, Goldstein MR, Peterson MJ, Tononi G, Benca RM (2011) Antidepressant effects of selective slow wave sleep deprivation in major depression: a high-density EEG investigation. J Psychiatr Res 45:1019–1026
10. Mohammadi M, Al-Azab F, Raahemi B, Richards G, Jaworska N, Smith D et al (2015) Data mining EEG signals in depression for their diagnostic value. BMC medical informatics and decision making 15:108
11. Acharya UR, Sudarshan VK, Adeli H, Santhosh J, Koh JE, Adeli A (2015) Computer-aided diagnosis of depression using EEG signals. Eur Neurol 73:329–336
12. Mohammed M, Khan MB, Bashier EBM. Machine learning: algorithms and applications: Crc press; 2016
13. Fatima M, Pasha M (2017) Survey of machine learning algorithms for disease diagnostic. J Intell Learn Syst Appl 9:1–16
14. Asri H, Mousannif H, Al Moatassime H, Noel T (2016) Using machine learning algorithms for breast cancer risk prediction and diagnosis. Procedia Computer Science 83:1064–1069
15. Ozcift A, Gulten A (2011) Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput Methods Prog Biomed 104:443–451
16. Palaniappan R, Sundaraj K, Sundaraj S (2014) A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals. BMC bioinformatics 15:223
17. Raghavendra U, Acharya UR, Adeli H (2019) Artificial intelligence techniques for automated diagnosis of neurological disorders. Eur Neurol 82:41–64
18. Jiang C, Li Y, Tang Y, Guan C (2021) Enhancing EEG-based classification of depression patients using spatial information. IEEE Transactions on Neural Systems and Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society:1
19. Sharma G, Parashar A, Joshi AM (2021) DepHNN: a novel hybrid neural network for electroencephalogram (EEG)-based screening of depression. Biomedical Signal Processing and Control 66:102393
20. Akbari H, Sadiq MT, Rehman AU (2021) Classification of normal and depressed EEG signals based on centered correntropy of rhythms in empirical wavelet transform domain. Health Information Science and Systems 9:1–15
21. Seal A, Bajpai R, Agnihotri J, Yazidi A, Herrera-Viedma E (2021) Krejcar O. A Deep Convolution Neural Networks Framework for Detecting Depression using EEG. IEEE Transactions on Instrumentation and Measurement, DeprNet
22. Kaur C, Bisht A, Singh P, Joshi G (2021) EEG signal denoising using hybrid approach of Variational mode decomposition and wavelets for depression. Biomedical Signal Processing and Control. 65:102337
23. Mitra V, Tsiartas A, Shriberg E. Noise and reverberation effects on depression detection from speech. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): IEEE; 2016. p. 5795–9
24. Afshan A, Guo J, Park SJ, Ravi V, Flint J, Alwan A. Effectiveness of Voice Quality Features in Detecting Depression. Interspeech2018. p. 1676–1680
25. Williamson JR, Quatieri TF, Helfer BS, Ciccarelli G, Mehta DD. Vocal and facial biomarkers of depression based on motor incoordination and timing. Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge2014. p. 65–72
26. Ooi KEB, Lech M, Allen NB (2012) Multichannel weighted speech classification system for prediction of major depression in adolescents. IEEE Trans Biomed Eng 60:497–506
27. Sturim D, Torres-Carrasquillo PA, Quatieri TF, Malyska N, Mc Cree A. Automatic detection of depression in speech using gaussian mixture modeling with factor analysis. Twelfth Annual Conference of the International Speech Communication Association2011
28. Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R, Nishimura M, Arai T (2018) Major depressive disorder discrimination using vocal acoustic features. J Affect Disord 225:214–220
29. Cohn JF, Kruez TS, Matthews I, Yang Y, Nguyen MH, Padilla MT, et al. Detecting depression from facial actions and vocal prosody. 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops: IEEE; 2009. p. 1–7
30. Mitra V, Shriberg E. Effects of feature type, learning algorithm and speaking style for depression detection from speech. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): IEEE; 2015. p. 4774–8
31. Williamson JR, Quatieri TF, Helfer BS, Horwitz R, Yu B, Mehta DD. Vocal biomarkers of depression based on motor incoordination. Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge2013. p. 41–8
32. Low L-SA, Maddage NC, Lech M, Sheeber LB, Allen NB (2010) Detection of clinical depression in adolescents' speech during family interactions. IEEE Trans Biomed Eng 58:574–586
33. Seneviratne N, Espy-Wilson C. Deep Learning Based Generalized Models for Depression Classification. arXiv preprint arXiv:201106739. 2020
34. Zhang L (2020) Duvvuri R. Nguyen T, Ghomi RH. Automated voice biomarkers for depression symptoms using an online cross-sectional data collection initiative. Depression and anxiety, Chandra KK
35. Dibeklioğlu H, Hammal Z, Cohn JF (2017) Dynamic multimodal measurement of depression severity using deep autoencoding. IEEE journal of biomedical and health informatics 22:525–536
36. Yildirim O, Talo M, Ciaccio EJ, San Tan R, Acharya UR (2020) Accurate deep neural network model to detect cardiac arrhythmia on more than 10,000 individual subject ECG records. Comput Methods Prog Biomed 197:105740
37. Soh DCK, Ng E, Jahmunah V, Oh SL, San Tan R, Acharya UR (2020) Automated diagnostic tool for hypertension using convolutional neural network. Comput Biol Med 126:103999
38. Panda R, Jain S, Tripathy R, Acharya UR (2020) Detection of shockable ventricular cardiac arrhythmias from ECG signals using

FFREWT filter-bank and deep convolutional neural network. Comput Biol Med 124:103939
39. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 103792
40. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition2016. p. 770–8
41. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition2017. p. 4700–8
42. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition2015. p. 1–9
43. Shensa MJ (1992) The discrete wavelet transform: wedding the a trous and Mallat algorithms. IEEE Trans Signal Process 40:2464–2482
44. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987
45. Vapnik V (1998) The support vector method of function estimation. Springer, Nonlinear Modeling, pp 55–85
46. Vapnik V. The nature of statistical learning theory: springer science & business media; 2013
47. Mumtaz W. MDD Patients and Healthy Controls EEG Data (New). figshare. Dataset. MDD Patients and Healthy Controls EEG Data generated by https://doi.org/10.6084/m9.figshare.4244171.v2. 2016
48. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. International Conference on Machine Learning: PMLR; 2017. p. 1263–1272
49. Ojala T, Pietikäinen M, Mäenpää T. A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. International Conference on Advances in Pattern Recognition: Springer; 2001. p. 399–408
50. Ahonen T, Hadid A, Pietikäinen M. Face recognition with local binary patterns. European conference on computer vision: Springer; 2004. p. 469–481
51. Liu L, Lao S, Fieguth PW, Guo Y, Wang X, Pietikäinen M (2016) Median robust extended local binary pattern for texture classification. IEEE Trans Image Process 25:1368–1381
52. Pan Z, Li Z, Fan H, Wu X (2017) Feature based local binary pattern for rotation invariant texture classification. Expert Syst Appl 88:238–248
53. Rafiee J, Tse P, Harifi A, Sadeghi M (2009) A novel technique for selecting mother wavelet function using an intelli gent fault diagnosis system. Expert Syst Appl 36:4862–4875
54. Avdakovic S, Nuhanovic A, Kusljugic M, Music M (2012) Wavelet transform applications in power system dynamics. Electr Power Syst Res 83:237–245
55. Goldberger J, Hinton GE, Roweis S, Salakhutdinov RR (2004) Neighbourhood components analysis. Adv Neural Inf Proces Syst 17:513–520
56. Kuncan F, Kaya Y, Kuncan M (2019) Sensör işaretlerinden cinsiyet tanıma için yerel ikili örüntüler tabanlı yeni yaklaşımlar. Journal of the Faculty of Engineering & Architecture of Gazi University 34
57. Kumar V, Minz S (2014) Feature selection: a literature review. SmartCR. 4:211–229
58. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Computers & Electrical Engineering 40:16–28
59. Tuncer T, Dogan S (2019) A novel octopus based Parkinson's disease and gender recognition method using vowels. Appl Acoust 155:75–83
60. Ezuma M, Erden F, Anjinappa CK, Ozdemir O, Guvenc I. Micro-UAV detection and classification from RF fingerprints using machine learning techniques. 2019 IEEE Aerospace Conference: IEEE; 2019. p. 1–13
61. Gao Y, Gao F (2010) Edited AdaBoost by weighted kNN. Neurocomputing. 73:3079–3088
62. Tuncer T, Dogan S, Pławiak P, Acharya UR (2019) Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. Knowl-Based Syst 186:104923
63. Bone D, Bishop SL, Black MP, Goodwin MS, Lord C, Narayanan SS (2016) Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. J Child Psychol Psychiatry 57:927–937
64. Mantri S, Patil D, Agrawal P, Wadhai V. Non invasive EEG signal processing framework for real time depression analysis. 2015 SAI Intelligent Systems Conference (IntelliSys): IEEE; 2015. p. 518–21
65. Acharya UR, Sudarshan VK, Adeli H, Santhosh J, Koh JE, Puthankatti SD et al (2015) A novel depression diagnosis index using nonlinear features in EEG signals. Eur Neurol 74:79–83
66. Erguzel TT, Sayar GH, Tarhan N (2016) Artificial intelligence approach to classify unipolar and bipolar depressive disorders. Neural Comput & Applic 27:1607–1616
67. Mumtaz W, Xia L, Mohd Yasin MA, Azhar Ali SS, Malik AS (2017) A wavelet-based technique to predict treatment outcome for major depressive disorder. PLoS One 12:e0171409
68. Liao S-C, Wu C-T, Huang H-C, Cheng W-T, Liu Y-H (2017) Major depression detection from EEG signals using kernel eigen-filter-bank common spatial patterns. Sensors. 17:1385
69. Kim AY, Jang EH, Kim S, Choi KW, Jeon HJ, Yu HY et al (2018) Automatic detection of major depressive disorder using electrodermal activity. Sci Rep 8:1–9
70. Cai H, Han J, Chen Y, Sha X, Wang Z, Hu B, Yang J, Feng L, Ding Z, Chen Y, Gutknecht J (2018) A pervasive approach to EEG-based depression detection. Complexity. 2018:1–13
71. Wu C-T, Dillon DG, Hsu H-C, Huang S, Barrick E, Liu Y-H (2018) Depression detection using relative EEG power induced by emotionally positive images and a conformal kernel support vector machine. Applied Sciences 8:1244
72. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H, Subha DP (2018) Automated EEG-based screening of depression using deep convolutional neural network. Comput Methods Prog Biomed 161:103–113
73. Sharma M, Achuth P, Deb D, Puthankattil SD, Acharya UR (2018) An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with EEG signals. Cogn Syst Res 52:508–520
74. Mumtaz W, Qayyum A (2019) A deep learning framework for automatic diagnosis of unipolar depression. Int J Med Inform 132:103983
75. Sandheep P, Vineeth S, Poulose M, Subha D. Performance analysis of deep learning CNN in classification of depression EEG signals. TENCON 2019–2019 IEEE Region 10 Conference (TENCON): IEEE; 2019. p. 1339–44
76. Li X, La R, Wang Y, Niu J, Zeng S, Sun S et al (2019) EEG-based mild depression recognition using convolutional neural network. Medical & biological engineering & computing 57:1341–1352
77. Mohammadi Y, Hajian M, Moradi MH. Discrimination of Depression Levels Using Machine Learning Methods on EEG Signals. 2019 27th Iranian Conference on Electrical Engineering (ICEE): IEEE; 2019. p. 1765–9
78. Ay B, Yildirim O, Talo M, Baloglu UB, Aydin G, Puthankattil SD, Acharya UR (2019) Automated depression detection using deep representation and sequence learning with EEG signals. J Med Syst 43:205
79. Duan L, Duan H, Qiao Y, Sha S, Qi S, Zhang X, Huang J, Huang X, Wang C (2020) Machine learning approaches for MDD detection and emotion decoding using EEG signals. Front Hum Neurosci 14

Emrah Aydemir Elazig in 1987. He received the M.S. degrees in computer teaching from the University of Elazig Firat, in 2012 and the Ph.D. degree in informatics from Istanbul University, Turkey, TR, in 2017. From 2012 to 2015, he was an Expert with the Istanbul Commerce University. From 2015 to 2021, he has been an Assistant Professor with the Computer Engineering Department, Kirsehir Ahi Evran University. Since 2021 he is Associate Professor in Sakarya University. He is the author of three books, more than 10 articles, and more than 40 conference presentation. His research interests include artificial intelligence, microcontroller, database and software.

Turker Tuncer received the master's degree in electronics and computer sciences and the Ph.D. degree in software engineering from Firat University, Elazig, Turkey, in 2011 and 2016, respectively. He is currently an Associate Professor with the Digital Forensics Engineering, Technology Faculty, Firat University. His main research interests include feature engineering, image processing, signal processing, information security, and pattern recognition. He has been working actively on developing algorithms in machine learning applied to visual surveillance and biomedical data.

Sengul Dogan received the master's degree in bioengineering the Ph.D. degree in electrical and electronics engineering from the Firat University, Elazig, Turkey, in 2007 and 2011, respectively. She is currently an Associate Professor with the Digital Forensics Engineering, Technology Faculty, Firat University. Her main research interests include computer forensics, mobile forensics, image processing, and signal processing. She has been working actively on developing algorithms in machine learning for biomedical data.

Raj Gururajan is a professor at University of Southern Queensland in School of Business. He has many science citiation index and international conference papers about machine learning and business systems. He has more than 250 papers with more than 1000 citations in Google Scholar (hindex=16).

U. R. Acharya PhD, DEng, DSc is a visiting faculty at Ngee Ann Polytechnic, Singapore. He is also (i) Adjunct Professor at University of Malaya, Malaysia, (ii) Adjunct Professor at Asia University, Taiwan, (iii) Associate faculty at Singapore University of Social Sciences, Singapore, (iv) Visiting Professor at Kumamoto University, Japan and (v) Adjunct Professor at University of Southern Australia, Australia. He received his Ph.D. from National Institute of Technology Karnataka (Surathkal, India), DEng from Chiba University (Japan) and DSc from AGH University of Science and Technology, Poland. He has published more than 500 papers, in refereed international SCI-IF journals (345), international conference proceedings (42), books (17) with more than 39,000 citations in Google Scholar (with h-index of 102). His world ranking is 224 and 4th in Singapore among top computer science scientists (https://www.guide2research.com/u/u-rajendra-acharya). He has worked on various funded projects, with grants worth more than 5.5 million SGD. He is ranked in the top 1% of the Highly Cited Researchers for the last five consecutive years (2016 to 2020) in Computer Science according to the Essential Science Indicators of Thomson. He owns five patents (3 obtained+ 2 filed) in application of artificial intelligence for healthcare applications. He is on the editorial board of many journals and has served as Guest Editor for many journals. His major academic interests are in Biomedical Signal Processing, Biomedical Imaging, Data mining, Visualization and Biophysics for better healthcare design, delivery and therapy. Please visit https://www.scholar.google.com.sg/citations?user=8FjY99sAAAAJ&hl=en for more details.