scientific reports

Check for updates

OPEN

Automated depression detection via cloud based EEG analysis with transfer learning and synchrosqueezed wavelet transform

Sara Bagherzadeh[1,2], Mohammad Reza Norouzi[2,3], Amirhesam Ghasri[2,4], Pouya Tolou Kouroshi[2,5], Sepideh Bahri Hampa[2,4], Fatemeh Farokhshad[1,2] & Ahmad Shalbaf [ID][6✉]

Post-COVID-19, depression rates have risen sharply, increasing the need for early diagnosis using electroencephalogram (EEG) and deep learning. To tackle this, we developed a cloud-based computer-aided depression diagnostic (CCADD) system that utilizes EEG signals from local databases. This system was optimized through a series of experiments to identify the most accurate model. The experiments employed a pre-trained convolutional neural network, ResNet18, fine-tuned on time–frequency synchrosqueezed wavelet transform (SSWT) images derived from EEG signals. Various data augmentation methods, including image processing techniques and noises, were applied to identify the best model for CCADD. To offer this device with minimal electrodes, we aimed to balance high accuracy with fewer electrodes. Two publicly databases were evaluated using this approach. Dataset I included 31 individuals detected with major depressive disorder and a control class of 27 age-matched healthy subjects. Dataset II comprised 90 participants, with 45 diagnosed with depression and 45 healthy controls. The leave-subjects-out cross-validation method with 20 subjects was used to validate the proposed method. The highest average accuracies for the selected model are 98%, 97%, 91%, and 88% for the parietal and central lobes in Databases I and II, respectively. The corresponding highest f-scores are 96.27%, 94.87%, 90.56%, and 89.65%. The highest intra-database accuracy and F1-score are 75.10% and 73.56% when training with SSWT images from Database II and testing with parietal images from Database I. This study introduces a novel cloud-based model for depression detection, paving the way for effective diagnostic tools and potentially revolutionizing depression management.

Keywords Convolutional neural network (CNN), Depression detection, Deep learning (DL), Electroencephalogram (EEG), Synchrosqueezing wavelet transform (SSWT)

Depression, a widespread mental disorder, is characterized by persistent feelings of sadness or diminished interest in activities over extended periods, as well as a sense of guilt, loss of concentration and energy, along with changes in sleep, appetite, and daily routines. Globally, an estimated 280 million people are affected by depression[1]. Depression can impact various aspects of life, including relationships with family, friends, and the community. Following the emergence of COVID-19, research indicates an increase in individuals experiencing depression[2]. Early diagnosis of depression can prevent the progression of the disease and its transformation into a more challenging and resistant condition[3].

[1]Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. [2]Mirolab Inc., Tehran, Iran. [3]Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran. [4]Department of Biological Sciences, Tarbiat Modares University, Tehran, Iran. [5]Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran. [6]Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ✉email: shalbaf@sbmu.ac.ir

nature portfolio

1

The identification of depression relies on a standardized questionnaire, which varies based on individual perspectives, contributing to potential misdiagnosis[4]. Various electrophysiological signals, including functional magnetic resonance imaging (fMRI)[5,6], magnetoencephalogram (MEG)[7,8], electroencephalogram (EEG)[9–12], and others, have been utilized. Among these various methods, EEG stands out due to its accessibility, cost-effectiveness, non-invasiveness, and high temporal resolution, making it the preferred diagnostic tool for identifying depression patients compared to alternatives like fMRI[13].

Related studies

Advanced computing techniques, such as machine learning and deep learning (DL), are utilized for the automated diagnosis of depression, leveraging EEG signals[14]. In some machine learning methods, the process involves steps like extracting features, selecting important ones, and then classifying the data. For example, Avots et al. investigated the use of machine learning methods for detecting and classifying depression[15]. By utilizing a dataset with EEG recordings from 10 healthy and 10 depressed individuals, the research incorporated linear and nonlinear features, such as relative band power, alpha power variability, and spectral asymmetry index. Various classifiers, including decision tree (DT), linear discriminant analysis (LDA) and support vector machine (SVM), were evaluated. The results demonstrated enhanced accuracy, with Relief-selected features boosting ensemble classifiers like AdaBoost, achieving an accuracy of 91.50%.

Recently, the preference for DL solutions has increased due to their capacity to learn and extract features from raw input data in automatic form, overcoming limitations faced by traditional machine learning methods and managing challenges in the manual analysis of EEG signals[16]. As a prominent DL approach, the Convolutional Neural Network (CNN) has become a significant and robust methodology for the development of computer-aided diagnosis systems. CNN models are utilized for EEG classification, involving the transformation of EEG data from signal space to image[17]. For instance, Acharya et al.[18] utilized CNN for EEG-based depression screening. Involving 15 healthy and 15 depressed subjects, the study employs a 10-fold cross-validation (CV) strategy on a test set, with accuracies for left and right hemisphere EEG signals averaging at 93.54% and 95.96%, respectively, over 10 iterations. In the pursuit of diagnosing major depressive disorder (MDD) through CNN, Loh et al.[19] proposed a DL model utilizing CNN for the automated detection of MDD through EEG signals. The model was trained and evaluated using spectrogram images derived from EEG signals via Short-Time Fourier Transform (STFT). Subsequently, these spectrogram images were fed into the CNN. EEG signals were sourced from a public database, consisting of 34 and 30 MDD and healthy subjects, respectively. The model employed 10-fold CV, achieving a classification accuracy of 99.58%. Xia et al.[20] presented a DL model to classify MDD patients using resting-state EEG data. Their proposed model integrates a multi-head self-attention mechanism among EEG channels, followed by a parallel two-branch CNN module to extract high-level features and a fully connected layer for classification. Additionally, three connectivity measures of transfer entropy (TE), phase locking value (PLV) and coherence (Coh) are used to compare results. The study utilized EEG data from 57 subjects, including 34 MDD patients and 30 healthy controls, and achieved an average classification accuracy of 91.06% using leave-one-subject-out cross-validation (LOSO CV). Saeedi et al.[21] proposed a DL framework for automatically discriminating between MDD patients and healthy controls using EEG signals. The study utilized generalized Partial Directed Coherence (gPDC) and Direct Directed Transfer Function (dDTF) methods to extract effective brain connectivity analysis. A combination of sixteen connectivity methods was employed to construct an image for each individual. In their study, which involved transforming EEG signals into connectivity-2D matrices, a 10-fold CV strategy was applied to a test dataset that included 30 healthy subjects and 34 depressed subjects, achieving an impressive accuracy of 99.24%. Khan et al.[22] introduced an automated diagnostic algorithm for Major Depressive Disorder (MDD) that leverages brain effective connectivity and a 3D CNN. The study utilized EEG signals from 30 individuals with MDD and 30 healthy control subjects to estimate the effective connectivity within the brain's default mode network (DMN). The effective connections among the six key regions of the DMN were then processed by the 3D CNN, resulting in a classification accuracy, sensitivity, and specificity of 100% for distinguishing between MDD and healthy control subjects. A study by Xia et al.[23] aimed to develop a deep spatial CNN (DSNet) for detecting MDD using EEG signals. The research utilized a public dataset comprising 30 MDD patients and 28 healthy controls (HCs) for evaluation. The methodology involved preprocessing the EEG signals, dividing them into five frequency bands, and implementing the DSNet model. The DSNet architecture comprises two convolutional layers, one pooling layer, and one fully connected (FC) layer, specifically designed to exploit spatial topological differences between MDD patients and healthy controls (HCs). The results demonstrated that DSNet achieved an impressive average classification accuracy of 91.69%, employing the LOSO CV strategy. Hashempour et al.[24] introduced a hybrid CNN-TCN model for continuous estimation of BDI scores. Their model achieved an average mean square error (MSE) of 5.64 (std = 1.6) and 9.53 (std = 2.94) for eyes open and closed conditions on 46 and 73 depressed and normal participants, respectively.

As discussed in previous articles[20–22], creating 2D images for CNNs often involves using the connectivity of different brain regions. However, unlike these traditional methods, our approach focuses on single channels. To design a wearable CCADD system, we aim to minimize the number of electrodes while maintaining diagnostic performance. Pre-trained CNNs are ideal for this purpose due to their high performance. Since CNNs require images as input, we need an imaging technique for single-channel EEG signals. Therefore, the connectivity measures mentioned in previous studies[20–22] are not suitable for our needs. Instead, time–frequency (TF) methods such as STFT, Fourier synchrosqueezed transform (FST), continuous Wavelet transform (CWT), discrete Wavelet transform (DWT) and synchrosqueezing Wavelet transform (SSWT) have been widely used in neuroscience studies for EEG signals[17,19,25–33]. TF methods are essential because EEG signals are non-stationary and time-varying. These methods (e.g., SSWT, CWT, STFT) transform EEG signals into 2D representations that capture both temporal and spectral information. This allows CNNs to analyze dynamic brain activity patterns associated with depression, such as abnormalities in delta, theta, or gamma bands, which are critical for accurate

diagnosis. In this study, we prefer SSWT due to its favorable properties. Also, this helps increase the number of channels and avoids overfitting issues. In the literature, there is a notable emphasis on TF representation when integrated with CNNs. Bagherzadeh et al.[25,26] utilized CWT to convert EEGs into two-dimensional color images. These images were subsequently processed by CNNs for recognition of emotion using EEGs. The method achieved the accuracy of over 95% in recognizing four classes of emotion from EEG signals. Shahabi et al.[17] utilized CWT in conjunction with CNN for classifying responders and non-responders to antidepressant treatment based on EEG signals. The study yielded significant results, which achieving recall, accuracy and specificity of 95.56%, 95.74% and 95.64%, respectively.

Nevertheless, these TF representations come with certain limitations. Fourier transform and wavelet-based TF representations require an experimental selection of window and kind of mother wavelet, respectively[26]. The synchrosqueezing transform (SST), which relies on the reassignment of coefficients within the TF matrix, is gaining attention for its effectiveness in TF analysis, particularly for producing sharper TF estimates compared to conventional methods[34,35]. Madhavan et al.[27] present a novel method to classify two classes of focal and non-focal from EEG signals, with the primary objective of automating the detection of focal epilepsy. The method utilizes Fourier and SST and a deep CNN for the analysis, aiming to accurately classify EEG signals without the need for feature extraction. The study evaluates the method using EEG signals from a public database and demonstrates superior performance in terms of specificity, accuracy and recall for the classification of mentioned classes. Bagherzadeh et al.[36] addressed challenges in EEG-based emotion recognition by proposing a portable Brain-Computer Interface (BCI) that utilizes two channels. Their approach integrates SSWT for enhanced TF analysis and ResNet-18 for automatic feature extraction. Emphasizing a user-friendly design, the system aims for practical real-world applications while minimizing computational costs. Four series of SEED database were employed for validation, demonstrating promising results with an average accuracy of 77.75% across diverse subjects.

Several key challenges arise in EEG-based depression detection. First, limited dataset sizes increase the risk of overfitting, often requiring mitigation through data augmentation and transfer learning techniques. Second, substantial inter-subject variability in EEG signals demands models with strong generalization capabilities. Third, traditional approaches relying on manually engineered features may fail to capture the complex, nonlinear patterns present in neural data. Additionally, computational complexity presents a significant hurdle, particularly when processing high-resolution EEG data for real-time or wearable applications. Finally, EEG signals are inherently susceptible to various noise sources and artifacts, necessitating robust preprocessing pipelines and advanced time–frequency analysis methods such as the SSWT.

Contribution of this study is as below:

- Develop a cloud-based computer-aided depression diagnostic (CCADD) system using EEG signals. Two publicly available depression databases are processed in two situations of inter-database and intra-database to demonstrate the possibility of CCADD. In this study two public databases from[37,38] are used to evaluate our proposed method.
- Effective TF representation from single EEG channel using SSWT. SSWT is advantageous for its ability to accurately represent non-stationary and time-varying signals, making it highly suitable for capturing the dynamic nature of EEG data in depression assessment.
- Improve diagnosis of depression through transfer learning (TL) technique. TL plays a crucial role in EEG-based depression detection for several key reasons. First, it enhances data efficiency by enabling the adaptation of pre-trained models (e.g., ResNet-18) to EEG analysis, significantly reducing the dependency on large labeled datasets. Second, it helps mitigate overfitting by leveraging robust feature representations learned from diverse source domains (e.g., ImageNet), thereby improving model generalization to new EEG data. The ResNet-18 model, being one of the most effective CNN models in terms of performance and quickness, is used in this study[39,40].
- A Comprehensive experiment using various data augmentation techniques to find most effective CNN model to accurately diagnose depression in inter-database and intra-databases situations through leave-subject-out CV (LSO CV).

Material and methods
Database I
In this study, we utilized the publicly available dataset from Mumtaz et al.[37]. The cohort included 34 patients (17 female) diagnosed with MDD, with a mean age of 40.3 years (SD = 12.9), and 30 age-matched healthy controls (9 female), with a mean age of 38.3 years (SD = 15.6). MDD diagnosis was confirmed using the Diagnostic and Statistical Manual-IV (DSM-IV) criteria[41]. To mitigate medication effects, a 2-week washout period was implemented prior to the initial EEG recording. The study protocol was approved by the Human Ethics Committee of Hospital Universiti Sains Malaysia (HUSM), Kelantan, Malaysia.

EEG data were acquired using a 19-electrode cap with electro-gel sensors arranged according to the international 10–20 system (electrodes: Fp1, F3, F7, Fz, Fp2, F4, F8, C3, C4, Cz, P3, Pz, P4, O1, O2, T3, T4, T5, T6). Signals were amplified using a Brain Master Discovery amplifier and sampled at 256 Hz. A 0.1–70 Hz band-pass filter and a 50 Hz notch filter were applied to minimize noise.

Five-minute recordings were obtained under both eyes-closed (EC) and eyes-open (EO) conditions, with participants seated in a relaxed state. For this study, only EO-condition EEG signals were processed. Data from three subjects per group were excluded due to technical errors. Detailed dataset characteristics are provided in Table 1.

Database	Sample size (total/MDD/control)	Subjects: Female (MDD/control), Male (MDD/control)	Age (m ± std)	BDI for MDD (m ± std)
Database I	64 (34/30)	26 (17/9), 38 (17/21)	40.3 ± 12.9	20.6 (± 8.6)
Database II	121 (46/75)	75 (34/41), 46 (12/34)	21 ± 2.3	22.22 (4.90)

Table 1. Details of both depression databases.

Dataset II

This study employed an openly accessible EEG dataset from OpenNeuro[38] consisting of 121 participants, including 46 individuals with depression (34 female) as determined by Beck Depression Inventory (BDI) scores exceeding 13 and 75 healthy controls. The University of Arizona granted ethical approval for the study, and all participants provided written informed consent. Cognitive assessment was performed using both the BDI and Spielberger Trait Anxiety Inventory (STAI) during mass and laboratory evaluations.

EEG recordings were obtained using a 64-channel Synamps2 system with Ag/AgCl electrodes, configured with a 0.5–100 Hz band-pass filter and 500 Hz sampling rate while maintaining impedances below 10 kΩ. Each participant completed two recording sessions (pre- and post-task), though only the pre-task resting-state EEG data were analyzed to prevent potential task-induced artifacts. The resting-state recording comprised 6 min of data collection, including 1-min intervals for both eyes-closed and eyes-open conditions.

For our time–frequency analysis, we selected 32 EEG channels (Af3, Af4, Fp1, F3, F7, Fz, Fp2, Ft7, Ft8, F4, F8, C3, C4, Cz, Cp1, Cp2, Cp3, Cp4, Cp5, Cp6, Cpz, P3, Pz, P4, O1, O2, T3, T4, T7, T8, Tp7, Tp8) from the extended 10–20 system. This channel reduction was implemented because our analysis did not require high spatial resolution, while still providing sufficient data for effective ResNet model optimization given our adequate sample size. Complete dataset specifications are provided in Table 1.

Proposed method

Figure 1 presents the workflow of our proposed depression detection method, which was evaluated using two distinct datasets (described in Sections "Database I" and "Dataset II"). During preprocessing, the EEG data were segmented into 30-s epochs, a duration selected after systematic evaluation of various window lengths (10, 15, and 60 s). This 30-s segmentation optimally balanced EEG pattern capture with classification performance while accommodating natural signal variability and providing adequate data for model fine-tuning.

The preprocessing pipeline included z-score normalization followed by downsampling to 128 Hz (from original sampling rates of 256 Hz for Data I and 500 Hz for Data II). We then transformed the preprocessed signals into time–frequency representations using the SSWT. These time–frequency images were analyzed using transfer learning with a pre-trained ResNet-18 architecture. Finally, we assessed model performance through LSO cross-validation in both intra-database and inter-database configurations.

Create TF image using SSWT

The synchrosqueezing transform is an innovative technique used to represent non-stationary EEG signals through TF method[35]. Daubechies et al. introduced an adaptive SSWT as a signal-processing method. Likewise, the wavelet transform disassembles signals by matching them with a set of TF elements. However, the analysis from the wavelet transform might be less sharp due to the associated TF spreading. The SSWT enhances TF representation sharpness and localization properties by synchro-squeezing and reallocating the TF spectrum[36]. It represents a post-processing technique following CWT. The CWT of the input EEG signal $x(t)$, utilizing $\psi(t)$ as the mother wavelet, and incorporating parameters c for scaling and b for translation, is expressed as:

$$W_x(c,d) = \frac{1}{\left|\sqrt{c}\right|} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-b}{c}\right)dt \tag{1}$$

In the context of SSWT, the TF coefficients $x(c,d)$ derived from Eq. (1) are squeezed into the instantaneous frequency trajectory. This process brings the representation closer to the ideal TF representation, which has high resolution in frequency, transforming $(c,d) \rightarrow (c, \omega_x(c,d))$ through the squeezing process. The SSWT can be explained in simpler terms by utilizing a set of wavelet coefficients derived from EEG signals, represented by Eq. (1):

$$T_x(\omega_n, d) = \sum_{a_l:|\omega_x(c_l,d)-\omega_n|\leq\frac{\Delta\omega}{2}} W_x(c_l, d) c^{-3/2}(\Delta c)_l \tag{2}$$

Here, ω_n denotes the frequency bins defined with a $\Delta\omega$ resolution, while a_l refers to the wavelet coefficients corresponding to specific frequency bands. In this study, we use the Bump wavelet as the mother wavelet $\psi(t)$, which is particularly effective for localized time–frequency analysis due to its smoothness and compact support in both time and frequency domains[36]. The EEG signals' SSWT can be reconstructed through this process:

$$x(d) = \Re\left[Z_\psi^{-1}\sum_n T_x(\omega_n, d)\Delta\omega\right] \tag{3}$$
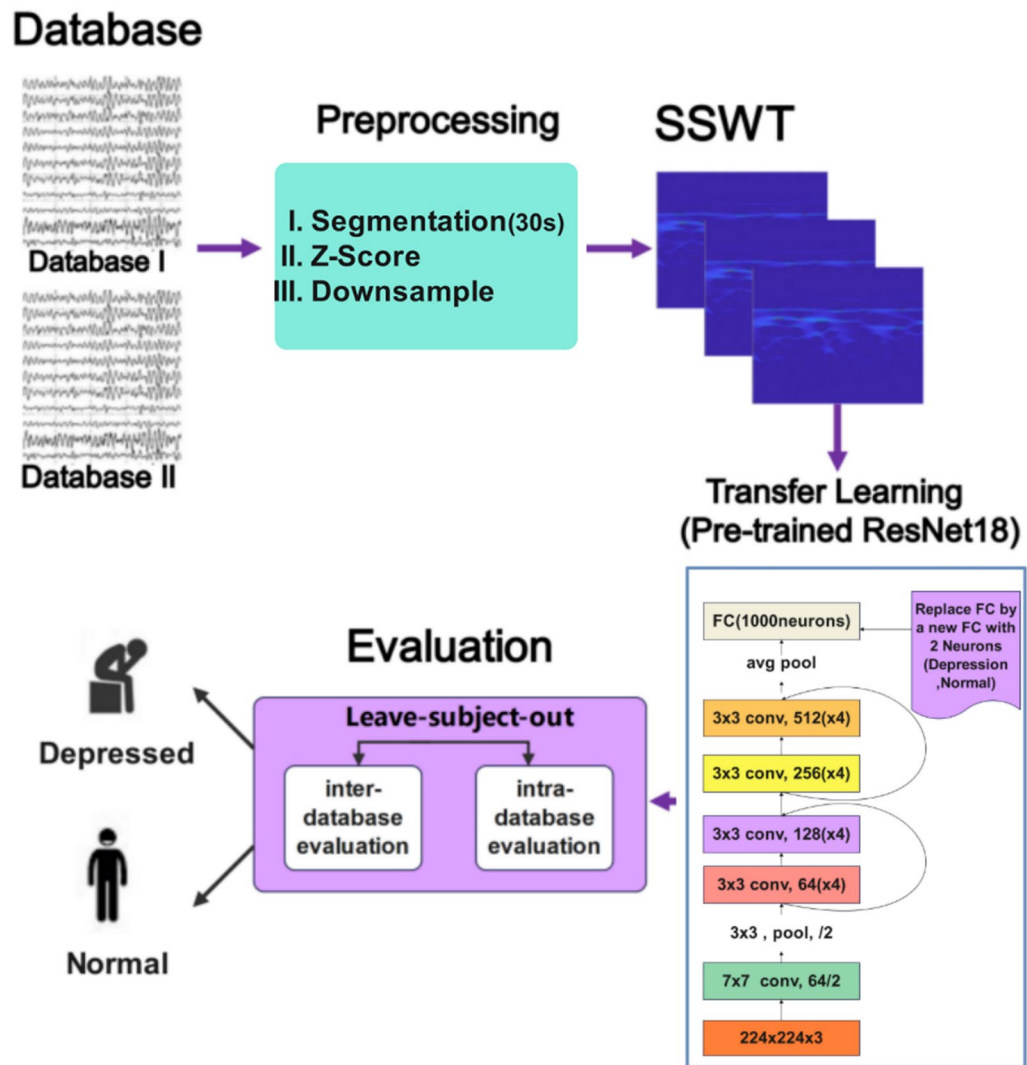
Fig. 1. Flowchart of the proposed depression detection system based on TF representation and ResNet-18 through both inter-database and intra-database strategies.

where $\mathfrak{R}$ returns the real component of the function contained within the square brackets, $Z_\psi = 0.5 \int_0^\infty \psi^*(\tau)\frac{d\tau}{\tau}$ denotes the normalization constant, while $\psi^*(\tau)$ signifies the Fourier transform of the mother wavelet[27].

ResNet-18

CNNs are an end-to-end DL technique widely used for automated feature extraction and data classification across various domains[42]. Their high performance, TL capabilities, and robustness make them particularly valuable in biomedical engineering applications. Accessing large-scale medical datasets remains challenging, often forcing researchers to work with limited data and adapt pre-trained models to achieve satisfactory performance even with small training sets[26]. Residual networks (ResNets), first introduced by He et al.[43], address key limitations in deep learning through their innovative residual units. These units consist of two consecutive convolutional layers operating in parallel with an identity shortcut connection, fundamentally changing how neural networks learn and propagate information.

The residual architecture provides several critical advantages over traditional deep networks. First, it effectively mitigates the vanishing gradient problem—a common issue where gradients diminish during backpropagation through deep layers, hampering network training[44,45]. The skip connections allow gradients to flow directly through the identity mapping path, maintaining signal strength across layers. Second, while enabling deeper architectures, residual units paradoxically reduce overall network complexity through implicit parameter sharing. Rather than learning complete transformations, the network focuses on learning residual functions, leading to more efficient parameter utilization. Third, the architecture enhances feature representation by facilitating both low-level and high-level feature capture, while enabling easy feature reuse from earlier layers through the additive skip connections.

Within the medical imaging domain, ResNet variants have demonstrated particular success in transfer learning applications. Among these, ResNet-18 has emerged as a preferred architecture[46], combining 8

residual units (comprising 17 convolutional layers) with a final fully-connected layer. As illustrated in Fig. 2, the architecture employs distinct processing blocks: convolutional layers (dark green), batch normalization/ReLU operations (light green), max pooling (orange), and fully-connected (FC)/classification layers (grey). This balanced design offers an optimal combination of depth and efficiency for medical image analysis tasks.

The evaluation technique

In this study, owing to a large statistical population, we employed a modified version of LOSO CV in two scenarios of inter-database and intra-database. In inter-database scenario, from a total of 58 and 90 subjects from Database I and II, SSWT images from 38 and 70 subjects were allocated for re-training, and 20 were retained for testing, respectively. Figure 3 illustrates this splitting method. Notably, we implemented LOS CV within these 20 subjects. By adopting this approach, a substantial number of individuals are retained for re-training, while the remainder are designated for testing. One notable advantage of this method is that it facilitates the development of a more generalized model, as a larger statistical population is utilized for re-training purposes. Also, the test process in both scenarios include two steps, in first step single channels are investigated in each step and in the
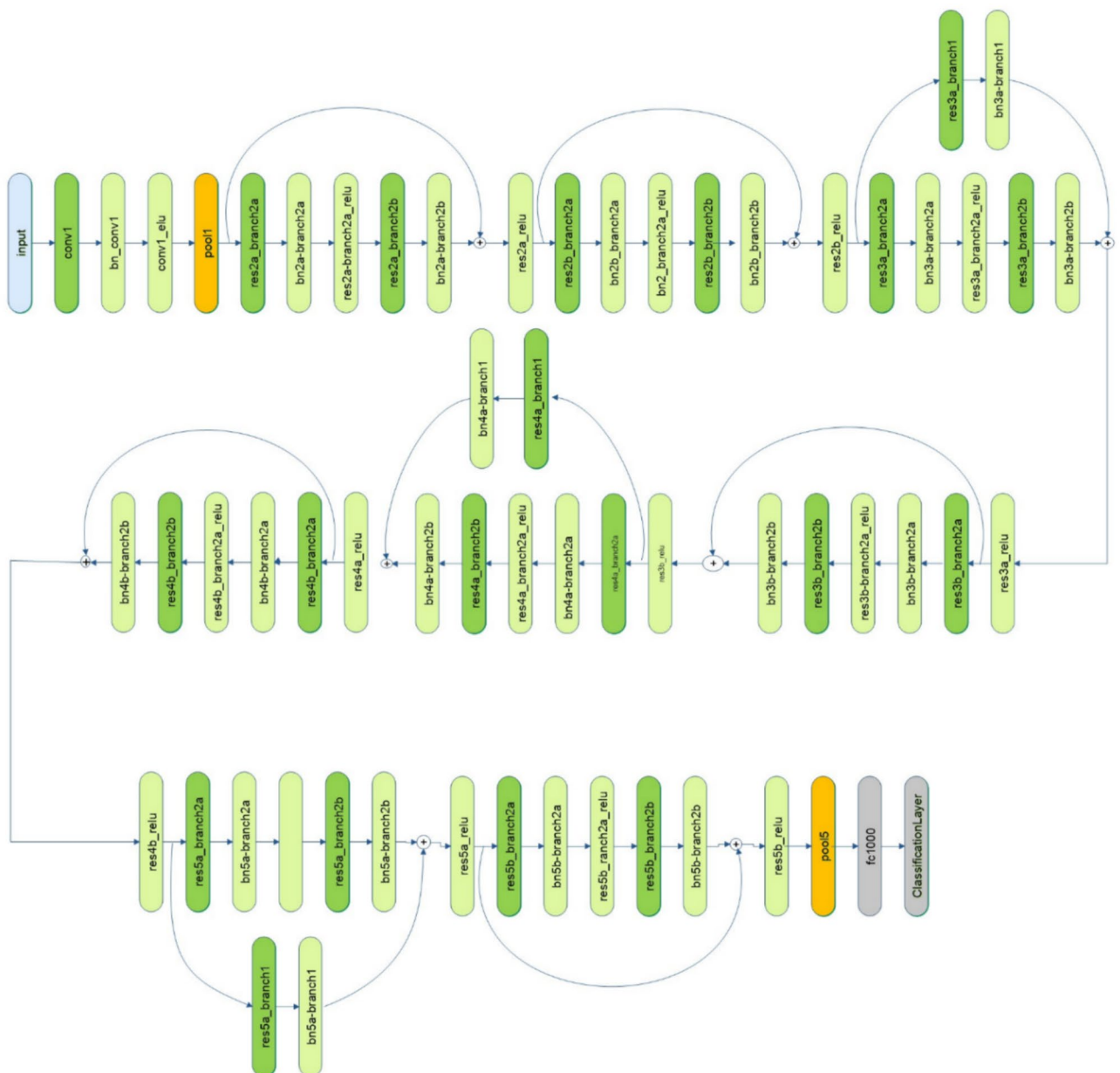
Fig. 2. depicts the Simplified structure of ResNet-18. The dark green blocks represent convolutional layers, light green blocks are batch normalization or ReLU, orange blocks signify max pooling with layer and the grey blocks are FC layer and classification layer. Additionally, each residual layer includes two parallel convolutional layers and an identity shortcut.
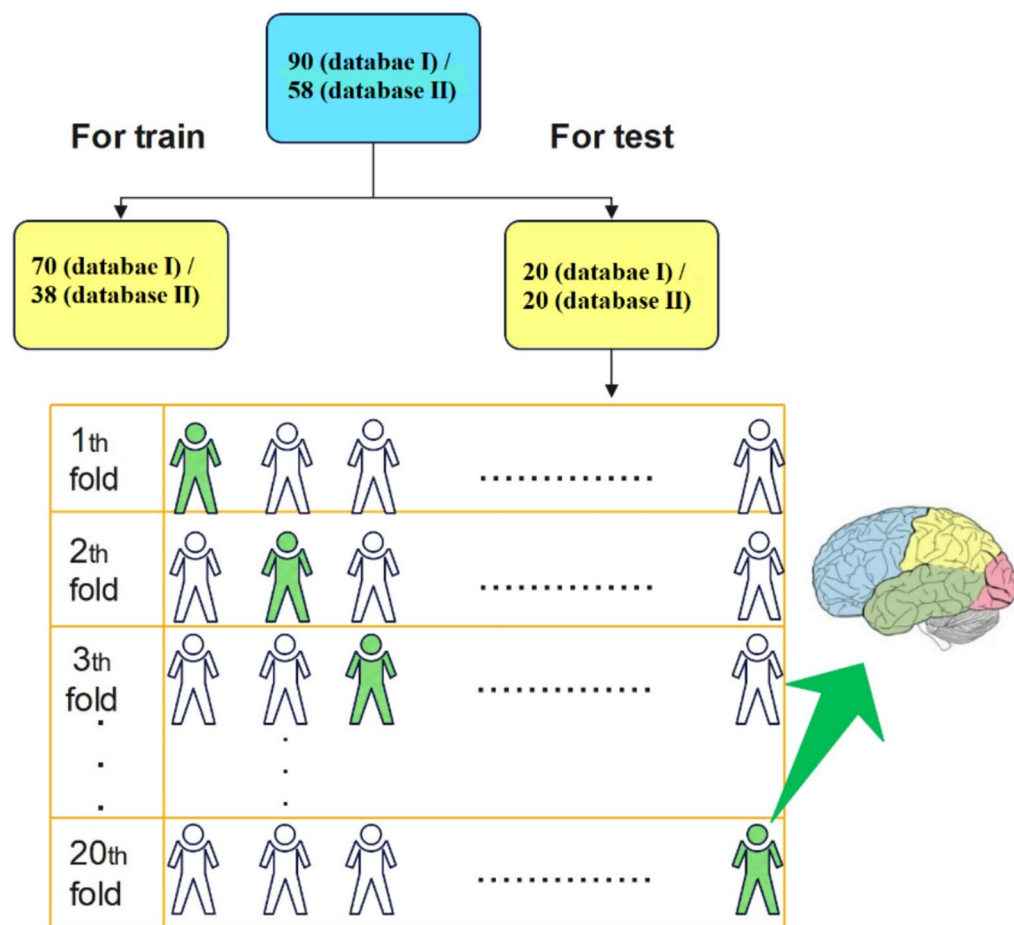
Fig. 3. Spilt subjects based on LSO method to evaluate proposed approach.

Lobe	Database I	Database II
Frontal	Fp1, F3, F7, Fz, Fp2, F4, F8	Af3, Af4, Fp1, F3, F7, Fz, Fp2, Ft7, Ft8, F4, F8
Central	C3, C4, Cz	C3, C4, Cz, Cp1, Cp2, Cp3, Cp4, Cp5, Cp6, Cpz
Parietal	P3, Pz, P4	P3, Pz, P4
Temporal	T3, T4, T5, T6	Tp7, Tp8, T3, T4, T5, T6, Tp7, Tp8
Occipital	O1, O2	O1, O2

Table 2. Separation of five brain lobes for each database.

second five brain lobes as mentioned in Table 2. Then, accuracy and F1-score are calculated to assess the model's performance on the test dataset. These metrics are defined as Eqs. (4) and (5)[47]:

$$\text{Accuracy (acc)} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$\text{F1} - \text{score} = \frac{2TP}{2TP + FN + FP} \tag{5}$$

In this context, TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

In the intra-database scenario, all SSWT images from Database I (and Database II) are used to retrain the desired CNN. In the two steps mentioned above, single channels and the five brain lobes, as outlined in Table 2, from all subjects of Database II (and Database I) are analyzed.

Results

Figure 4 illustrates one 30-s segment of raw EEG signal from one healthy subject and one depressed patient from Database I (a and b) and II (c and d). The vertical and horizontal axes are amplitude of signal and time (30 s). As mentioned in section "Dataset II", EEG signals from both databases were segmented using non-overlapping window with the length of 30 s. Each segment was then normalized using the Z-score method. EEG signals from Database I and II were downsampled from 250 and 500 to 128 Hz to reduce computational load, respectively.

Then, SSWT was calculated for each segment and then plotted in color scales and saved in jpg format. Figure 5a,c display SSWT representations of EEG signals from healthy subjects, while Fig. 5b,d show corresponding SSWT images for depressed subjects in Database I and Database II, respectively. Vertical and horizontal axes are frequency components and time points of each signal, respectively.

Finally, 40,864 and 49,216 SSWT images from 45 healthy subjects and 45 depressed subjects were created from 32 EEG channels (according to 10–20 electrode placement) from Database II. Also, 5130 and 4617 SSWT images (30 s segments) from depressed and normal subjects from 19 EEG channels of 31 depressed patients and 27 healthy subjects were created from Database I. According to our evaluation strategy, 10 depressed and 10 healthy subjects were considered as the test set for both databases. Then, ResNet-18 was fine-tuned using SSWT images from the train set. Hyperparameters of this CNN are listed in Table 3. Different data augmentation methods are examined to test performance of detection using our proposed method. These conditions are reported in Table 4.
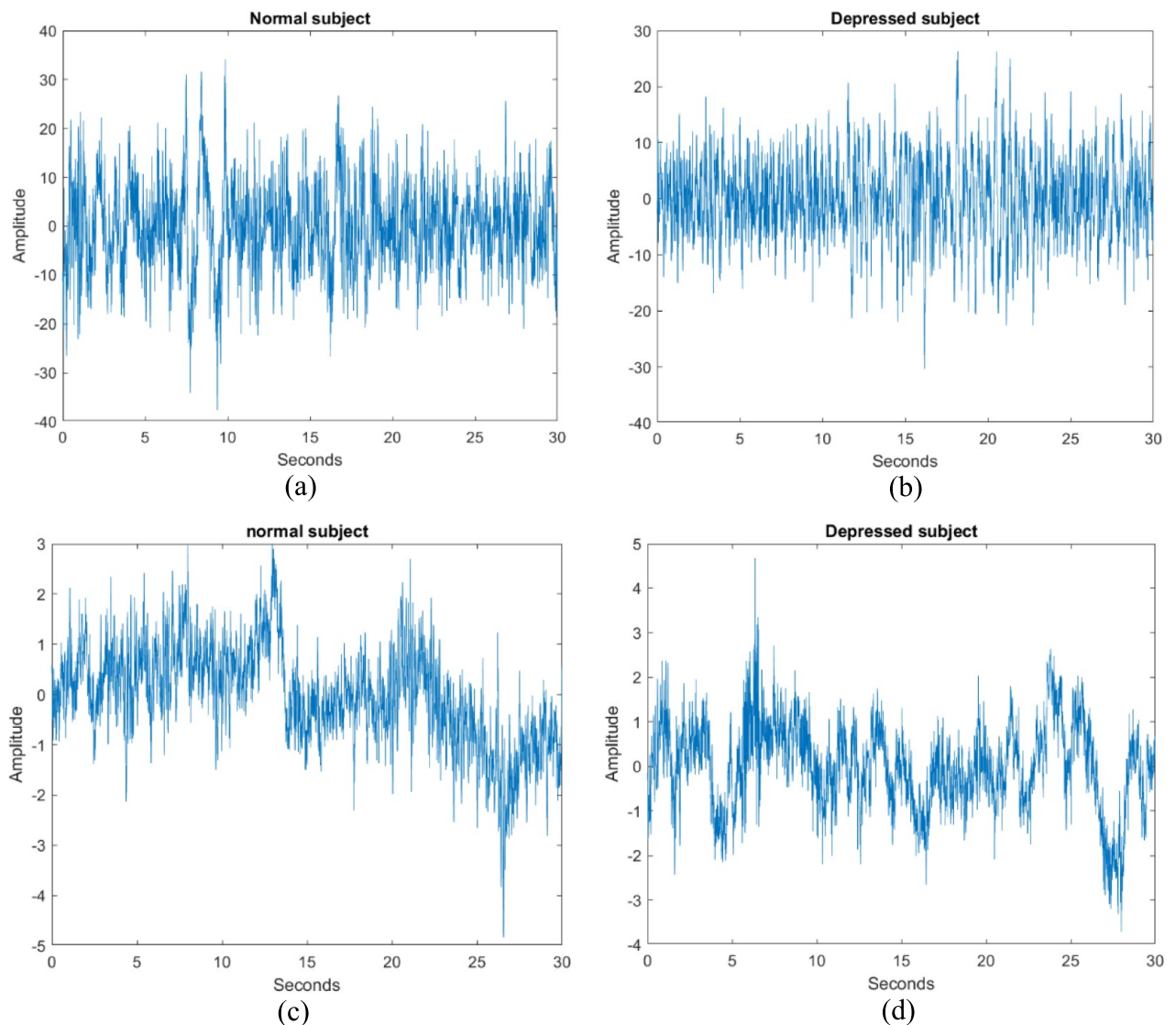
Fig. 4. Example EEG signal segments from (**a,b**) Database I and (**c,d**) Database II, each spanning 30 s. The vertical axis represents the signal amplitude (in microvolts, μV), and the horizontal axis represents time (in seconds). These segments illustrate raw EEG activity from different subject groups (normal vs. depressed), highlighting variations in brainwave patterns.
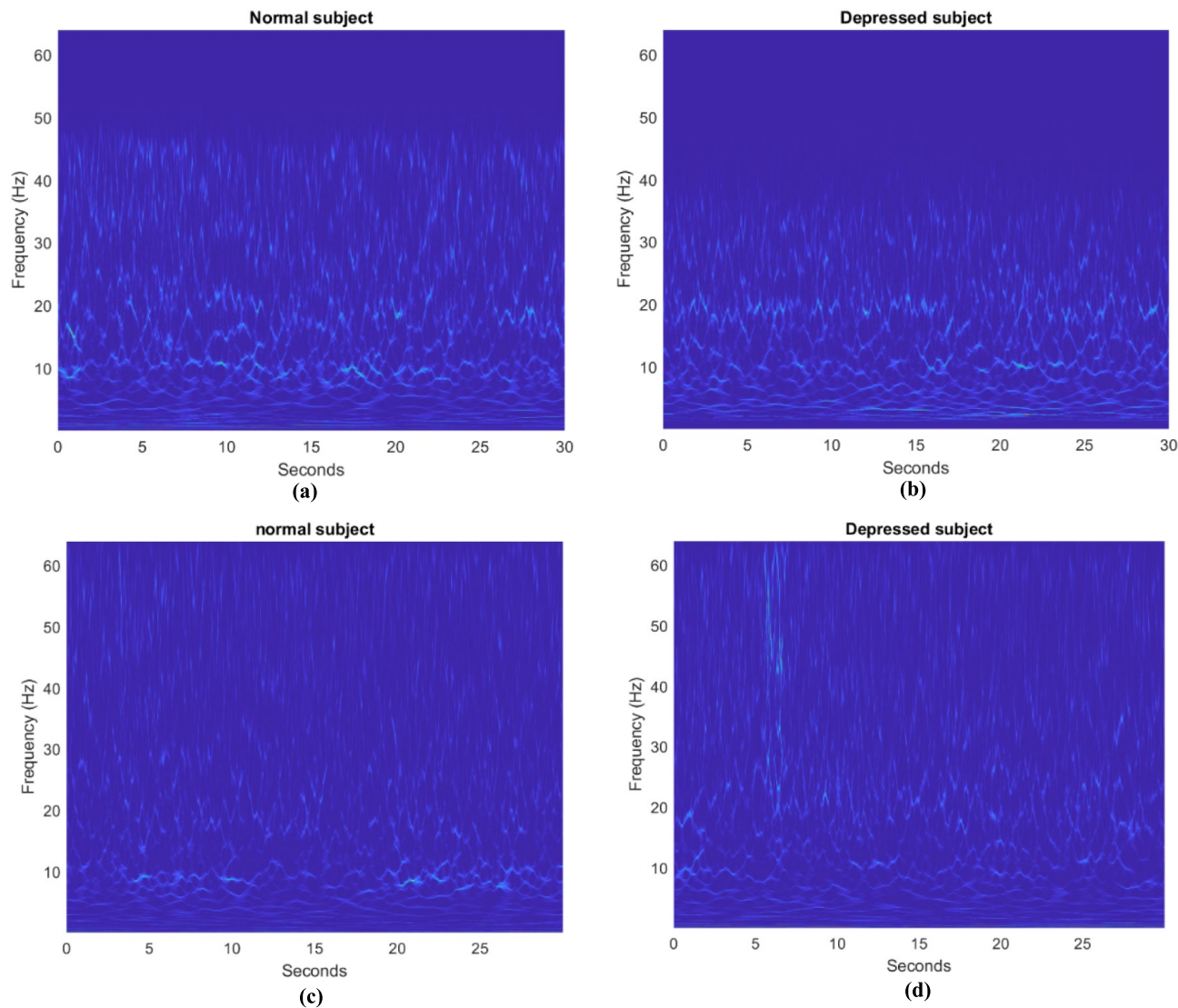
Fig. 5. Synchrosqueezed Wavelet Transform (SSWT) time–frequency representations of the EEG signals in Fig. 4: (**a,c**) normal subjects and (**b,d**) depressed subjects for Database I and II, respectively. The vertical axis shows frequency components (Hz), and the horizontal axis represents time (s). SSWT enhances the visualization of non-stationary EEG features, where depressive states may exhibit distinct energy distributions across frequency bands.

Mini batch	32	Squared gradient Decay factor	0.99
Initial learning rate	0.0008	Optimizer	ADAM
Epoch	30		

Table 3. ResNet-18 hyperparameters.

Inter-database situations

In this section, 9 mentioned models are fine-tuned on 44 and 70 subjects from Database I and II, respectively. Tables 5 and 6 report average accuracy values obtained from twenty subjects which were out using these models per channels based on colors, i.e., 19 and 32 EEG channels for Database I and II, respectively. In these tables, the green spectrum indicates accuracy values > 86% (average of channels over models) for Database I and > 75% for Database II and red spectrum represents lower values. According to Table 5, P4, P3, Pz, T3, T4, T5, T6, O1 and O2 channels achieved higher than 90% accuracies using all models especially Model 3. According to Table 6, C4, C3, T3, T4, F3, F4, Fp1, Fp2, F7, F8, Af3, Af4 and Cp1 channels achieved higher than 80% accuracies using all models especially model 3.

# Condition	Data augmentation technique	# Condition	Data augmentation technique
1	No	6	Random Rotation [− 30, 30]
2	Random shear X [− 20, 20]	7	Random reflection X, Random Rotation [− 30, 30]
3	Random reflection X	8	Random reflection X, Random scale X [2, 2]
4	Random reflection Y	9	Gaussian noise ($\sigma^2 = 0.05$)
5	Random scale X [2, 2]		

Table 4. Conditions of training process using ResNet18 and MobileNetV2 on SSWT images of both databases.

Channels	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fp1	77.00%	80.00%	82.00%	80.00%	78.00%	77.00%	82.00%	81.00%	78.00%
Fp2	80.00%	83.00%	86.00%	80.00%	83.00%	84.00%	82.00%	82.00%	81.00%
F3	76.00%	80.00%	82.00%	79.00%	80.00%	82.00%	80.00%	79.00%	89.40%
F4	82.00%	82.00%	86.00%	85.00%	84.00%	84.00%	83.00%	80.00%	82.00%
F7	77.00%	80.00%	82.00%	80.00%	82.00%	80.00%	78.00%	79.00%	78.00%
F8	78.00%	81.00%	84.00%	80.00%	81.00%	82.00%	83.00%	80.00%	79.00%
Fz	81.00%	83.00%	86.00%	82.00%	80.00%	84.00%	82.00%	84.00%	80.00%
C3	80.00%	84.00%	86.00%	83.00%	82.00%	80.00%	79.00%	85.00%	82.00%
C4	80.00%	82.00%	86.00%	84.00%	86.00%	85.00%	80.00%	82.00%	83.00%
Cz	82.00%	83.00%	86.00%	83.00%	81.00%	82.00%	80.00%	82.00%	84.00%
P4	88.00%	91.00%	94.00%	90.00%	87.00%	85.00%	92.00%	90.00%	91.00%
P3	87.00%	90.00%	94.00%	89.00%	88.00%	88.00%	90.00%	88.00%	89.00%
Pz	85.00%	87.00%	90.00%	86.00%	88.00%	83.00%	90.00%	86.00%	87.00%
T3	86.00%	88.00%	92.00%	82.00%	90.00%	88.00%	86.00%	88.00%	90.00%
T4	88.00%	90.00%	96.00%	86.00%	91.00%	92.00%	88.00%	90.00%	89.00%
T5	86.00%	91.00%	94.00%	84.00%	86.00%	82.00%	84.00%	89.00%	82.00%
T6	78.00%	85.00%	90.00%	82.00%	89.00%	90.00%	82.00%	80.00%	79.00%
O1	93.50%	94.00%	94.50%	94.00%	86.00%	92.00%	86.00%	91.00%	89.00%
O2	94.00%	93.00%	94.50%	93.00%	82.00%	90.00%	89.00%	93.00%	86.00%

Table 5. Accuracy values from each single channel, averaged over all subjects from Database I using the nine models listed in Table 4, in an inter-database scenario (trained on Database I and tested on Database I). The green spectrum signifies higher accuracy values (above 86% on average across channels and models), whereas the red spectrum signifies lower accuracy values.

Then, SSWT of channels according to five brain lobes (Table 2) are collected to feed into these 9 models, for example, frontal includes Af3, Af4, Fp1, F3, F7, Fz, Fp2, Ft7, Ft8, F4, F8 channels. Figures 6 and 7 show accuracy (a) and F1-score (b) values obtained from averaging on twenty subjects which were out using mentioned models for these brain lobes for Database I and II, respectively. According to Figs. 6 and 7, the third model, ResNet-18 while augment with random reflection in direction X, achieves the highest accuracies and F-scores among test subjects. According to Fig. 6, the three highest average accuracies achieved using this model are 98%, 97% and 95% for parietal, central and frontal lobes. Additionally, the three highest average F-scores for these lobes achieved 96.27%, 94.87% and 94.36%, respectively. According to Fig. 7 the three highest average accuracies for parietal, central and frontal lobes of Database II achieved 91%, 88% and 86%, respectively. Also, the three highest average accuracies for these lobes of Database II achieved 90.56%, 89.65% and 84.37%, respectively.

| Channels	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fp1	0.77	0.80	0.83	0.80	0.81	0.80	0.79	0.80	0.79
Fp2	0.76	0.78	0.81	0.77	0.78	0.78	0.79	0.78	0.77
F3	0.78	0.82	0.84	0.83	0.82	0.82	0.81	0.80	0.79
F4	0.75	0.77	0.82	0.78	0.76	0.77	0.77	0.78	0.76
F7	0.74	0.78	0.81	0.79	0.79	0.80	0.78	0.80	0.77
F8	0.78	0.80	0.82	0.80	0.81	0.80	0.8	0.78	0.78
Fz	0.69	0.72	0.75	0.74	0.70	0.72	0.72	0.71	0.7
C3	0.77	0.8	0.83	0.8	0.81	0.82	0.8	0.81	0.79
C4	0.78	0.83	0.85	0.82	0.84	0.82	0.83	0.83	0.81
Cz	0.69	0.74	0.76	0.73	0.74	0.72	0.74	0.73	0.72
P4	0.68	0.70	0.73	0.7	0.72	0.7	0.71	0.72	0.67
P3	0.69	0.72	0.75	0.72	0.73	0.72	0.71	0.72	0.71
Pz	0.68	0.70	0.72	0.70	0.70	0.71	0.70	0.70	0.68
T3	0.77	0.78	0.82	0.79	0.80	0.79	0.79	0.8	0.78
T4	0.78	0.8	0.83	0.8	0.80	0.80	0.78	0.79	0.78
T5	0.68	0.73	0.74	0.72	0.72	0.73	0.7	0.72	0.7
T6	0.68	0.72	0.74	0.7	0.72	0.71	0.72	0.7	0.69
O1	0.7	0.73	0.74	0.7	0.72	0.72	0.72	0.7	0.71
O2	0.68	0.72	0.74	0.73	0.73	0.72	0.74	0.71	0.7
Af3	0.78	0.8	0.82	0.81	0.8	0.82	0.81	0.8	0.79
Af4	0.77	0.79	0.8	0.78	0.8	0.79	0.79	0.8	0.78
Cp1	0.77	0.78	0.8	0.79	0.78	0.78	0.79	0.8	0.77
Cp2	0.69	0.73	0.74	0.73	0.72	0.72	0.73	0.72	0.71
Cp3	0.7	0.74	0.76	0.73	0.75	0.74	0.73	0.74	0.72
Cp4	0.71	0.74	0.76	0.75	0.72	0.74	0.72	0.73	0.73
Cp5	0.71	0.72	0.76	0.73	0.76	0.75	0.74	0.74	0.72
Cp6	0.69	0.72	0.74	0.71	0.72	0.69	0.72	0.69	0.68
Cpz	0.7	0.73	0.76	0.73	0.74	0.74	0.75	0.72	0.69
Ft7	0.69	0.74	0.74	0.73	0.73	0.71	0.74	0.73	0.71
Ft8	0.73	0.76	0.78	0.75	0.76	0.75	0.72	0.73	0.74
Tp7	0.71	0.73	0.76	0.75	0.74	0.72	0.71	0.74	0.71
Tp8	0.69	0.74	0.74	0.7	0.73	0.72	0.73	0.72	0.68

Table 6. Accuracy values from each single channel, averaged over all subjects from Database II using the nine models listed in Table 4, in an inter-database scenario (trained on Database II and tested on Database II). The green spectrum signifies higher accuracy values (above 75% on average across channels and models), whereas the red spectrum signifies lower accuracy values.

Intra-databases situation

In this section, as mentioned in section "The evaluation technique", we investigate the intra-database scenario in two steps, in first step, all SSWT images from each database is used as train set and each channel from the other database is used as test, separately. The second step investigate five brain lobes as before. Results of the first step (accuracy values) for each database is reported in Tables 7 and 8. Results of the second step (accuracy values)
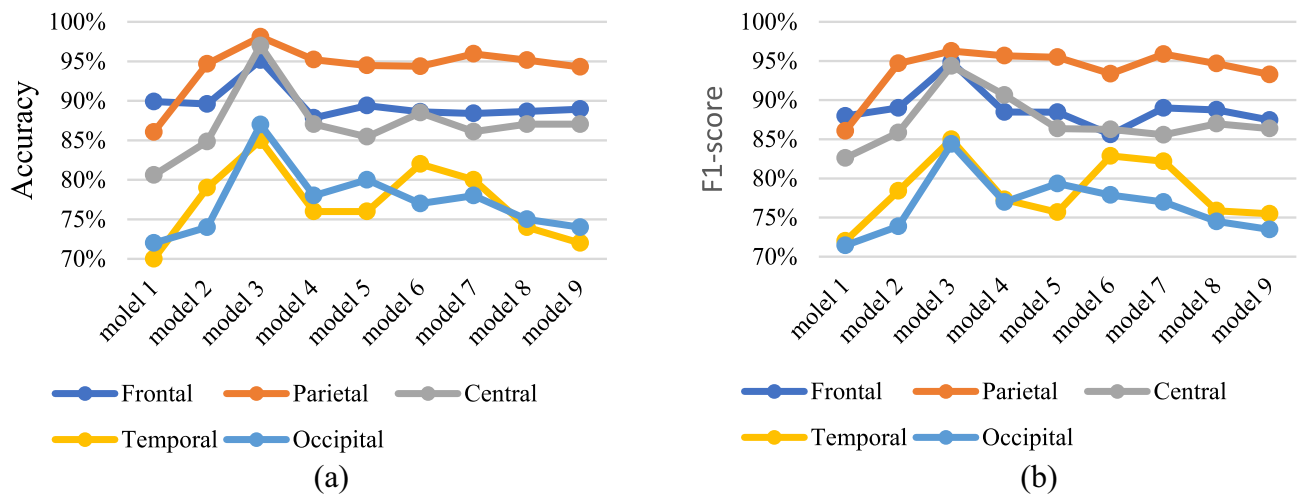
Fig. 6. Performance comparison of nine machine learning models for depression classification on Database I: (**a**) accuracy and (**b**) F1-score, averaged across 20 out-of-sample subjects. Results are stratified by five brain lobes (frontal, temporal, parietal, occipital, central) to assess regional discriminative power. Higher values indicate better model performance in distinguishing depressed versus normal EEG patterns.
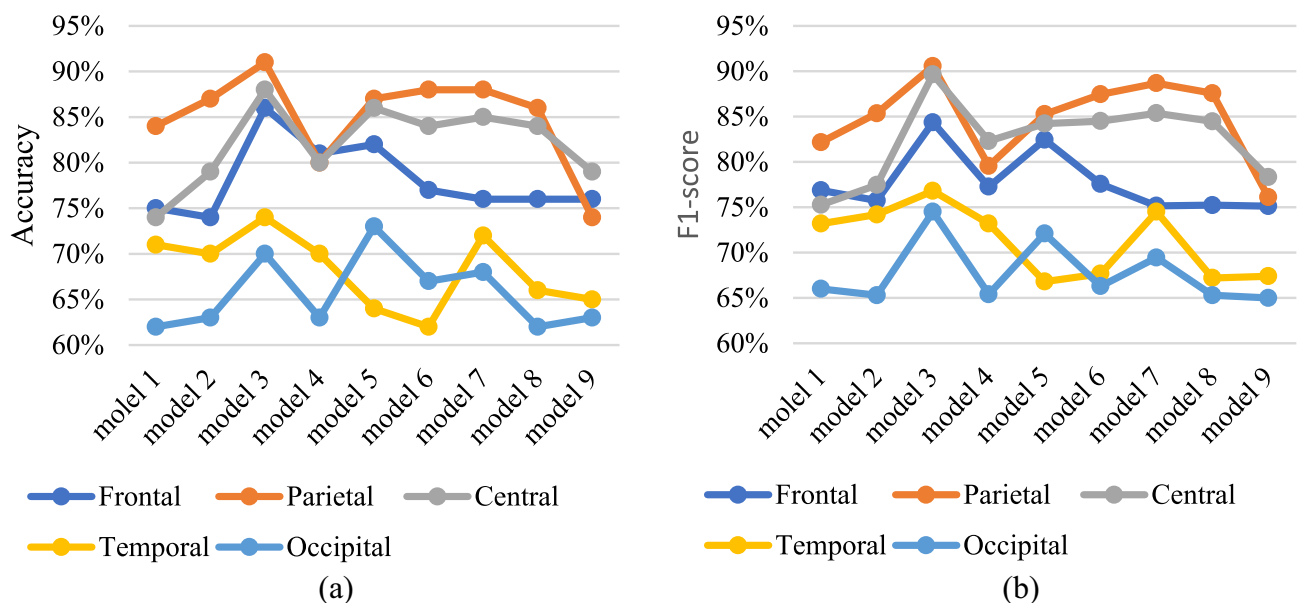
Fig. 7. Performance comparison of nine machine learning models for depression classification on Database II: (**a**) accuracy and (**b**) F1-score, averaged across 20 out-of-sample subjects. Results are stratified by five brain lobes (frontal, temporal, parietal, occipital, central) to assess regional discriminative power. Higher values indicate better model performance in distinguishing depressed versus normal EEG patterns.

is illustrated in Figs. 8 and 9 for Database I and II, respectively. The green spectrum signifies higher accuracy values (above 60% (64% for Database II) on average across channels and models), whereas the red spectrum signifies lower accuracy values. According to Table 7, Fp2, F8, Fz, C3, P3, P4, T3, T4, T5, T6, O1 and O2 channels achieved higher than 60% accuracies using at most three models especially Model 3. According to Table 8, most frontal channels, T3, T4, T6, Tp7, P3, P4, Cp1, Cp4 and Cp5 channels achieved higher than 70% accuracies using all models especially model 3.

Figures 8 and 9 present the results of intra-databases, i.e., training with Database I and test with SSWT images from separate five lobes from Database II and vice versa. The highest performance of detection of depression achieves form both databases using the parietal lobe. The highest accuracy and F1-score for these databases achieves 68.50%, 67.21%, 75.10 and 73.56%, respectively.

Column1	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fp1	60%	59%	60%	58%	60%	59%	58%	57%	58%
Fp2	58%	59%	62%	59%	60%	56%	59%	57%	58%
F3	59%	59%	60%	60%	60%	60%	59%	58%	59%
F4	58%	59%	61%	59%	58%	58%	58%	60%	60%
F7	58%	59%	60%	58%	59%	58%	57%	59%	57%
F8	60%	62%	62%	59%	60%	60%	58%	60%	60%
Fz	61%	60%	63%	58%	60%	58%	60%	60%	58%
C3	57%	62%	62%	58%	60%	58%	60%	59%	58%
C4	58%	60%	61%	58%	56%	56%	59%	60%	60%
Cz	57%	57%	59%	55%	58%	55%	58%	58%	58%
P4	60%	60%	61%	58%	59%	60%	60%	60%	58%
P3	58%	60%	63%	59%	61%	59%	59%	59%	60%
Pz	59%	57%	60%	55%	58%	57%	57%	58%	57%
T3	60%	57%	60%	60%	60%	57%	55%	56%	57%
T4	58%	60%	61%	59%	62%	59%	58%	56%	55%
T5	59%	60%	62%	57%	60%	58%	59%	56%	56%
T6	60%	58%	60%	56%	58%	59%	60%	59%	59%
O1	58%	56%	59%	58%	58%	58%	56%	55%	55%
O2	57%	59%	58%	56%	57%	55%	55%	56%	59%

Table 7. Accuracy values from each single channel, averaged over all subjects from Database II using the nine models listed in Table 4, in an intra-database scenario (trained on Database I and tested on Database II). The green spectrum signifies higher accuracy values (above 60% on average across channels and models), whereas the red spectrum signifies lower accuracy values.

Discussion

The aim of this study is to design a highly precise DL-CCAD system for diagnosing depression using a minimal number of EEG channels, suitable for integration into smart wearable devices such as headbands, headphones, or hands-free accessories. Given the wearable nature of the CCAD system, it is crucial to use fewer electrodes while maintaining diagnostic performance. EEG is essential in detecting and diagnosing depression, providing an objective basis for assessment. To enhance diagnostic accuracy, the study introduces a novel approach that integrates the SSWT and TL using the ResNet18 architecture. SSWT enhances traditional TF methods by providing sharper localization through the reassignment of TF coefficients, resulting in more concentrated representations that improve discriminability between depressed and healthy subjects. Additionally, SSWT effectively handles the non-stationary nature of EEG signals, which is critical for capturing depression-related dynamic patterns. Its compatibility with CNNs further enables the generation of high-resolution TF images suitable for deep learning models, facilitating automatic extraction of discriminative features.

To optimize depression classification, we employ TL with ResNet18, leveraging its pre-trained feature extraction capabilities to efficiently capture high-level representations of EEG features relevant to depression. Pre-trained CNNs like ResNet18 are ideal for this task due to their proven performance in image-based pattern recognition. Since CNNs require image inputs, we transform single-channel EEG signals using TF imaging techniques. While connectivity measures are commonly used in other studies[20–22], they are unsuitable here; instead, TF methods—such as STFT, CWT, and notably SSWT—are preferred due to their established efficacy in neuroscience applications[17,19,25,26]. Among these, SSWT is selected for its superior ability to handle the dynamic and non-stationary properties of EEG signals.

ResNet-18 was selected for its superior speed in fine-tuning and overall performance. It fine-tuned three to five times faster than EfficientNetB0, NasNetMobile, Xception, and Inception on our specified GPU within the Matlab environment. Additionally, its accuracy surpassed that of other quick pre-trained CNNs such as ShuffleNet, MobileNetV2, SqueezeNet, and DarkNet19. As illustrated in Figs. 6, 7, 8 and 9, Model 3 achieved the highest accuracy among all models. This model employed random reflection in the X direction to augment

Channels	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fp1	64.00%	67.00%	70.00%	67.00%	68.00%	67.00%	66.00%	67.00%	66.00%
Fp2	63.00%	65.00%	68.00%	64.00%	65.00%	65.00%	66.00%	65.00%	64.00%
F3	65.00%	69.00%	71.00%	70.00%	69.00%	69.00%	68.00%	67.00%	66.00%
F4	62.00%	64.00%	69.00%	65.00%	63.00%	64.00%	64.00%	65.00%	63.00%
F7	61.00%	65.00%	68.00%	66.00%	66.00%	67.00%	65.00%	67.00%	64.00%
F8	65.00%	67.00%	69.00%	67.00%	68.00%	67.00%	67.00%	65.00%	65.00%
Fz	56.00%	59.00%	62.00%	61.00%	57.00%	59.00%	59.00%	71.00%	57.00%
C3	64.00%	67.00%	70.00%	67.00%	68.00%	69.00%	67.00%	68.00%	66.00%
C4	65.00%	70.00%	72.00%	69.00%	71.00%	69.00%	70.00%	70.00%	68.00%
Cz	56.00%	61.00%	63.00%	60.00%	61.00%	59.00%	61.00%	60.00%	59.00%
P4	55.00%	57.00%	60.00%	57.00%	59.00%	57.00%	71.00%	59.00%	54.00%
P3	56.00%	59.00%	62.00%	59.00%	60.00%	59.00%	71.00%	59.00%	71.00%
Pz	55.00%	57.00%	59.00%	57.00%	57.00%	71.00%	57.00%	57.00%	55.00%
T3	64.00%	65.00%	69.00%	66.00%	67.00%	66.00%	66.00%	67.00%	65.00%
T4	65.00%	67.00%	70.00%	67.00%	67.00%	67.00%	65.00%	66.00%	65.00%
T5	55.00%	60.00%	61.00%	59.00%	59.00%	60.00%	57.00%	59.00%	57.00%
T6	55.00%	59.00%	61.00%	57.00%	59.00%	71.00%	59.00%	57.00%	56.00%
O1	57.00%	60.00%	61.00%	57.00%	59.00%	59.00%	59.00%	57.00%	71.00%
O2	55.00%	59.00%	61.00%	60.00%	60.00%	59.00%	61.00%	71.00%	57.00%
Af3	65.00%	67.00%	69.00%	68.00%	67.00%	69.00%	68.00%	67.00%	66.00%
Af4	64.00%	66.00%	67.00%	65.00%	67.00%	66.00%	66.00%	67.00%	65.00%
Cp1	64.00%	65.00%	67.00%	66.00%	65.00%	65.00%	66.00%	67.00%	64.00%
Cp2	56.00%	60.00%	61.00%	60.00%	59.00%	59.00%	60.00%	59.00%	71.00%
Cp3	57.00%	61.00%	63.00%	60.00%	62.00%	61.00%	60.00%	61.00%	59.00%
Cp4	71.00%	61.00%	63.00%	62.00%	59.00%	61.00%	59.00%	60.00%	60.00%
Cp5	71.00%	59.00%	63.00%	60.00%	63.00%	62.00%	61.00%	61.00%	59.00%
Cp6	56.00%	59.00%	61.00%	71.00%	59.00%	56.00%	59.00%	56.00%	55.00%
Cpz	57.00%	60.00%	63.00%	60.00%	61.00%	61.00%	62.00%	59.00%	56.00%
Ft7	56.00%	61.00%	61.00%	60.00%	60.00%	71.00%	61.00%	60.00%	71.00%
Ft8	60.00%	63.00%	65.00%	62.00%	63.00%	62.00%	59.00%	60.00%	61.00%
Tp7	71.00%	60.00%	63.00%	62.00%	61.00%	59.00%	71.00%	61.00%	71.00%
Tp8	56.00%	61.00%	61.00%	57.00%	60.00%	59.00%	60.00%	59.00%	55.00%

Table 8. Accuracy values from each single channel, averaged over all subjects from Database I using the nine models listed in Table 4, in an intra-database scenario (trained on Database II and tested on Database I). The green spectrum signifies higher accuracy values (above 64% on average across channels and models), whereas the red spectrum signifies lower accuracy values.

(a)

(b)
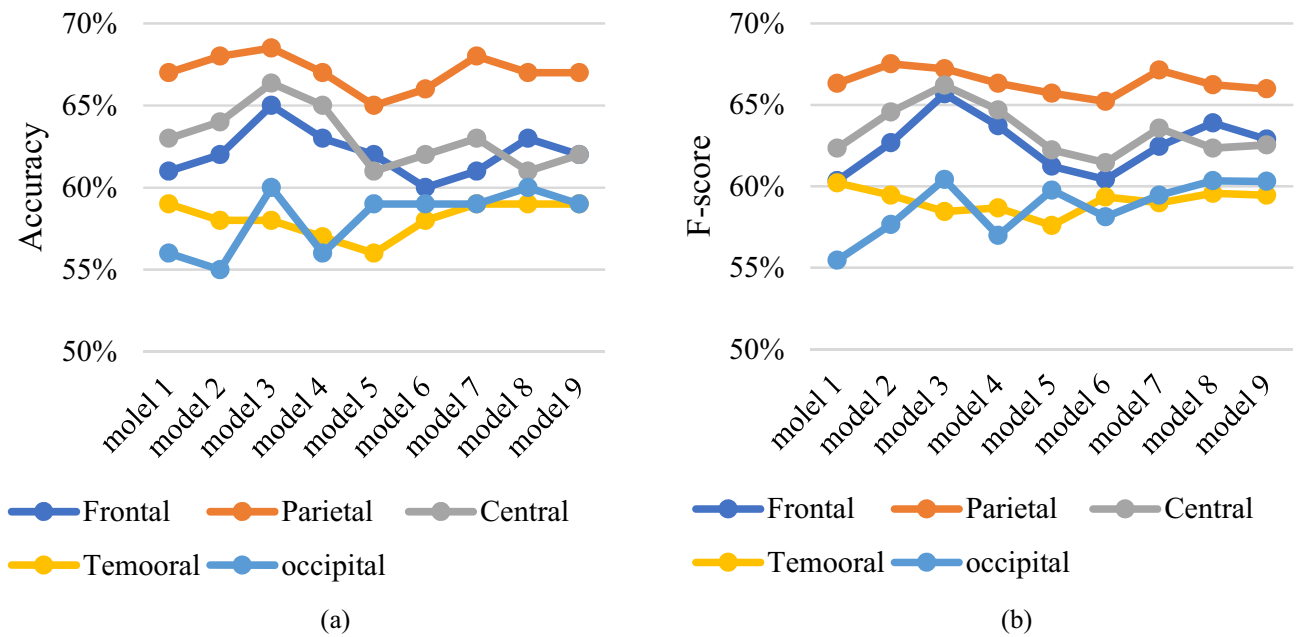
Fig. 8. Intra-database evaluation (training on Database I and testing based on five brain lobes) of depression detection using all EEG channels: (**a**) accuracy and (**b**) F1-score for nine models, tested on brain lobes. This assesses model generalizability between datasets.

(a)

(b)
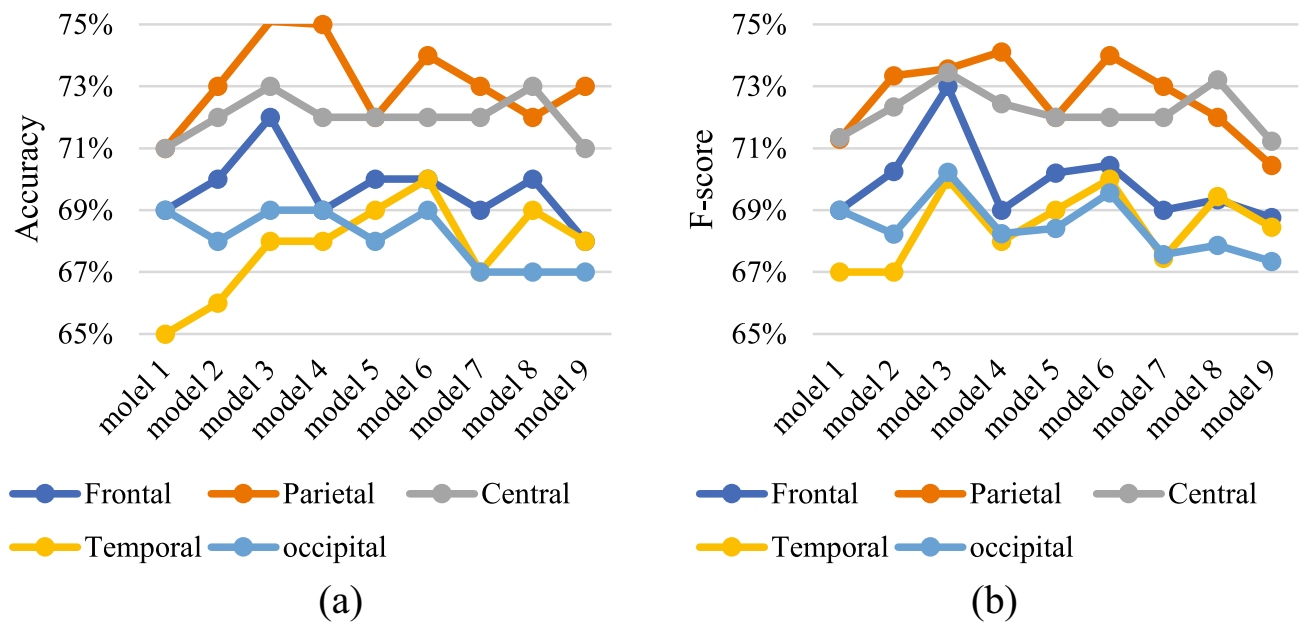
Fig. 9. Intra-database evaluation (training on Database II and testing based on five brain lobes) of depression detection using all EEG channels: (**a**) accuracy and (**b**) F1-score for nine models, tested on brain lobes. This assesses model generalizability between datasets.

SSWT images in the training set. This augmentation likely benefits from the invariant nature of patterns in both depression and normal classes, aiding in detection under these conditions. However, in other conditions, operations like shear appear to alter the patterns of both classes (depression and normal), thereby reducing detection performance.

Although our findings do not separate standard frequency bands, Fig. 5b,d show that depressed and normal subjects differ significantly in the delta and gamma frequency bands. This observation aligns with several studies[20,48–50]. Fitzgerald and Watson[48] declare the gamma oscillations as biomarker of MDD subjects. De Aguair

Study (year)	Database	Processing methods	Evaluation	Accuracy
[13] (2020)	I	Frequency-dependent multilayer brain, multilayer deep CNN	10-fold CV	97.27%
[22] (2023)	I	DSNet	LOSO CV	91.69%
[20] (2023)	I	Multi-head self-attention, connectivity measures (TE, PLV, Coh), CNN	LOSO CV	91.06%
[10] (2022)	I	10 selected EEG data, Inception time model	10-fold CV	90.10%
[4] (2023)	I	Paired asymmetry and DFA (temporal, parietal), SVM	Hold out CV	89.24%
[53] (2022)	I	WCOH, 2D CNN	10-fold CV	98.10%
Ours (2024)	I, II	SSWT (parietal lobe), data augmentation, ResNet-18	LSO CV	98% (Database I), 91% (Database II)

Table 9. Comparison of our results with recent DL studies on same databases. Articles on database II are almost based on classification of multiclass of depression and could not compare with this study.

Neto et al.[49] highlighted the theta and gamma frequency bands as promising for depression detection. Nobakhsh et al.[50] identified delta and theta bands as discriminative using the dDTF effective connectivity measure.

In our study, as shown in Tables 5, 6, 7 and 8 multiple frontal, temporal, central and parietal single-channels have been obtained high accuracies in both scenarios. For example, Fp2 achieved the accuracy of 86% and 81% from Database I and II, respectively. This channel is similar with result of Nobakhsh al.[50], which reported the Fp2 channel as a biomarker for depression detection. Moreover, according to Tables 7 and 8, this channel achieved the accuracy of 62% and 68% in intra-databases scenario, respectively. In general, to design a low-cost, single-channel wearable CCADD system with medium accuracy, the single channels represented in dark green in Tables 7 and 8 are suitable.

According to Figs. 6, 7, 8 and 9, we found that the parietal, central, and frontal lobes significantly distinguish between MDD (or depressed people) and the control group, consistent with findings in[4,50]. Khadidos et al.[4] achieved the best results using paired asymmetry and DFA features from the temporal and parietal lobes. Thibodeau et al.[51] and Shahabi et al.[52] also identified the frontal lobe as the most discriminative for depressed patients. In general, to design a comfortable, medium-cost, and more accurate wearable CCADD system, just the parietal lobe (central or frontal) is suitable. The results of this study demonstrated that the parietal lobe has a high capability in distinguishing individuals with depression from the control group. This finding highlights the importance of this region in analyzing EEG signals related to depression. Given the superior performance of this lobe, it is feasible to design low-cost and accurate wearable systems for depression diagnosis using a limited number of EEG channels. Such a system could be easily implemented in the form of headbands or other wearable devices. We demonstrate that training on Database II cause the accuracy and F1-score of 75.10% and 73.56%, respectively. Also, comparing Figs. 8 and 9, we found that, the performance of training all nine models on Database II is higher than on Database I. This improvement could be attributed to the larger number of samples (created images) and subjects in Database II, which increased the models' generalizability and flexibility.

Compare with state-of-the-art studies

Table 9 presents a comparative analysis of the findings from our study with recent research on depression recognition. In the pursuit of accurate depression detection using EEG data, researchers have explored various methodologies to optimize performance. We acknowledge that direct comparison of results across studies in Table 9 requires caution due to differences in evaluation protocols (e.g., 10-fold CV, LOSO CV, LSO CV) and dataset usage (e.g., single-database vs. cross-database validation). For instance, while[53] reports 98.10% accuracy with 10-fold CV on Database I, our LSO CV protocol provides a more rigorous estimate of generalizability by ensuring completely independent test subjects. Additionally, studies on Database II often focus on multiclass depression classification, limiting direct comparability with our binary task. These methodological variations highlight the need for standardized evaluation frameworks in future work.

Advantages and limitations of the study

Advantages of our work are:

- Achieving a high level of model generalization by incorporating SSWT images, TL, and data augmentation in our proposed algorithm.
- Obtaining good results in both the comparative inter-database and intra-database approaches highlights the generalizability of our model
- Identifying the distinctive region in depression detection, which is the parietal lobe, through the utilization of SSWT images and TL.

Limitations of our work are:

However, it's worth noting some limitations of our study. Firstly, the size and diversity of the datasets might impact the general applicability of our findings to wider populations. Additionally, the choice of specific parameters and hyperparameters in our methodology may impact the performance of the classification model.

Conclusion

In this study, our objective was to present a comprehensive cloud-based model for the detection of depression. Leveraging the robustness of DL, we utilized SSWT, a relatively new image input, which demonstrated a

significant advancement over conventional CWT methods. By incorporating these novel SSWT images into various data augmentation methods, we systematically evaluated nine top-performing models. The results of these models on the test population revealed the parietal lobe (98% for Dataset I, 91% for Dataset II), central lobe (97% for Dataset I, 88% for Dataset II), and frontal lobe (95% for Dataset I, 86% for Dataset II) as the most distinctive areas differentiating between depressed and normal individuals. Given the significance of our findings and the prominence of the mentioned lobes, it is evident that our results could contribute to the design and development of a headset capable of detecting depression. Additionally, our results prove the effectiveness of the proposed model on unseen databases, achieving an average accuracy and F1-score of 75.10% and 73.56% across both Datasets I and II.

In the future, we aim to improve our detection methods to increase intra-database results. We also plan to develop a CCADD system to estimate levels of depression, helping to prevent participants from developing the disorder. Moreover, we will adapt the wearable device for routine use, enhancing its applicability in everyday life.

Data availability

In this study we have used two open access databases: Dataset I: https://figshare.com/articles/dataset/EEG_Data_New/4244171. Dataset II: https://openneuro.org/datasets/ds003478/versions/1.1.0.

References

1. World Health Organization (WHO) (2017) Depression and other common mental disorders: Global health estimates. World Health Organization. https://apps.who.int/iris/handle/10665/25461
2. Gabarrell-Pascuet, A. et al. The association of social support and loneliness with symptoms of depression, anxiety, and posttraumatic stress during the COVID-19 pandemic: A meta-analysis. *Int. J. Environ. Res. Public Health*. https://doi.org/10.3390/ijerph20042765 (2023).
3. Davey, C. G. & McGorry, P. D. Early intervention for depression in young people: A blind spot in mental health care. *Lancet Psychiatry*. https://doi.org/10.1016/S2215-0366(18)30292-X (2019).
4. Khadidos, A. O. et al. Computer aided detection of major depressive disorder (MDD) using electroencephalogram signals. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3262930 (2023).
5. Mousavian, M., Chen, J. & Greening, S. Depression detection using atlas from fMRI images. In *Proceedings of 19th IEEE International Conference* (2020). https://doi.org/10.1109/ICMLA51294.2020.00210
6. Mousavian, M. et al. Depression detection from sMRI and rs-fMRI images using machine learning. *J. Intell. Inf. Syst.* https://doi.org/10.1007/s10844-021-00653-w (2021).
7. Lu, Q. et al. Multichannel matching pursuit of MEG signals for discriminative oscillation pattern detection in depression. *Int. J. Psychophysiol.* https://doi.org/10.1016/j.ijpsycho.2013.04.010 (2013).
8. Kähkönen, S. et al. Dysfunction in early auditory processing in major depressive disorder revealed by combined MEG and EEG. *J. Psychiatry Neurosci.* **32**(5), 316–322 (2007).
9. Sarkar, A., Singh, A. & Chakraborty, R. A deep learning-based comparative study to track mental depression from EEG data. *Neurosci. Inform.* **1**, 1. https://doi.org/10.1016/j.neuri.2022.100039 (2022).
10. Rafiei, A. et al. Automated detection of major depressive disorder with EEG signals: A time series classification using deep learning. *IEEE Access*. https://doi.org/10.1109/ACCESS.2022.3190502 (2022).
11. Afzali, A., Khaleghi, A., Hatef, B., Akbari Movahed, R. & Pirzad Jahromi, G. Automated major depressive disorder diagnosis using a dual-input deep learning model and image generation from EEG signals. *Waves Random Complex Media*. https://doi.org/10.1080/17455030.2023.2187237 (2023).
12. Mirjebreili, S. M., Shalbaf, R. & Shalbaf, A. Prediction of treatment response in major depressive disorder using a hybrid of convolutional recurrent deep neural networks and effective connectivity based on EEG signal. *Phys. Eng. Sci. Med.* https://doi.org/10.1007/s13246-024-01392-2 (2024).
13. Dang, W. et al. Multilayer brain network combined with deep convolutional neural network for detecting major depressive disorder. *Nonlinear Dyn.* https://doi.org/10.1007/s11071-020-05665-9 (2020).
14. Squires, M. et al. Deep learning and machine learning in psychiatry: A survey of current progress in depression detection, diagnosis and treatment. *Brain Inform.* https://doi.org/10.1186/s40708-023-00188-6 (2023).
15. Avots, E. et al. Ensemble approach for detection of depression using EEG features. *Entropy*. https://doi.org/10.3390/e24020211 (2022).
16. Safayari, A. & Bolhasani, H. Depression diagnosis by deep learning using EEG signals: A systematic review. *Med. Nov. Technol. Devices* **1**, 1. https://doi.org/10.1016/j.medntd.2021.100102 (2021).
17. Sadat Shahabi, M., Shalbaf, A. & Maghsoudi, A. Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG. *Biocybern Biomed. Eng.* https://doi.org/10.1016/j.bbe.2021.06.006 (2021).
18. Acharya, U. R. et al. Automated EEG-based screening of depression using deep convolutional neural network. *Comput. Methods Programs Biomed.* https://doi.org/10.1016/j.cmpb.2018.04.012 (2018).
19. Loh, H. W. et al. Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals. *Expert Syst.* https://doi.org/10.1111/exsy.12773 (2022).
20. Xia, M. et al. An end-to-end deep learning model for EEG-based major depressive disorder classification. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3270426 (2023).
21. Saeedi, A. et al. Major depressive disorder diagnosis based on effective connectivity in EEG signals: A convolutional neural network and long short-term memory approach. *Cogn. Neurodyn.* https://doi.org/10.1007/s11571-020-09619-0 (2021).
22. Khan, D. M. et al. Automated diagnosis of major depressive disorder using brain effective connectivity and 3D convolutional neural network. *IEEE Access*. https://doi.org/10.1109/ACCESS.2021.3049427 (2021).
23. Xia, M. et al. *DSNet: EEG-Based Spatial Convolutional Neural Network for Detecting Major Depressive Disorder* (Springer, 2023). https://doi.org/10.1007/978-981-19-8222-4_5.
24. Hashempour, S. et al. Continuous scoring of depression from EEG signals via a hybrid of convolutional neural networks. *IEEE Trans. Neural Syst. Rehabil. Eng.* https://doi.org/10.1109/TNSRE.2022.3143162 (2022).
25. Bagherzadeh, S. et al. Emotion recognition using continuous wavelet transform and ensemble of convolutional neural networks through transfer learning from electroencephalogram signal. *Front. Biomed. Technol.* https://doi.org/10.18502/fbt.v10i1.11512 (2023).

26. Bagherzadeh, S. et al. A hybrid EEG-based emotion recognition approach using wavelet convolutional neural networks (WCNN) and support vector machine. *Basic Clin. Neurosci.* https://doi.org/10.32598/bcn.2021.3133.1 (2023).
27. Madhavan, S., Tripathy, R. K. & Pachori, R. B. Time-frequency domain deep convolutional neural network for the classification of focal and non-focal EEG signals. *IEEE Sens. J.* https://doi.org/10.1109/JSEN.2019.2956072 (2020).
28. Puri, D. V., Kachare, P. H. & Nalbalwar, S. L. Metaheuristic optimized time–frequency features for enhancing Alzheimer's disease identification. *Biomed. Signal Process. Control* **1**(94), 106244. https://doi.org/10.1016/j.bspc.2024.106244 (2024).
29. Puri, D. V., Gawande, J. P., Kachare, P. H. & Al-Shourbaji, I. Optimal time-frequency localized wavelet filters for identification of Alzheimer's disease from EEG signals. *Cogn. Neurodyn.* **19**(1), 12. https://doi.org/10.1007/s11571-024-10198-7 (2025).
30. Puri, D. et al. Hybrid reptile-snake optimizer based channel selection for enhancing Alzheimer's disease detection. *J. Bionic Eng.* **29**, 1–7. https://doi.org/10.1007/s42235-024-00636-x (2025).
31. Gorur, K., Olmez, E., Ozer, Z. & Cetin, O. EEG-driven biometric authentication for investigation of Fourier synchrosqueezed transform-ICA robust framework. *Arab. J. Sci. Eng.* **48**(8), 10901–10923 (2023).
32. Ozturk, H., Eraslan, B. & Gorur, K. Investigation of t-SNE and dynamic time warping within a unified framework for resting-state and minor analysis visual task-related EEG alpha frequency in biometric authentication: A detailed analysis. *Digital Signal Process.* **1**, 105042 (2025).
33. Gorur, K. & Eraslan, B. The single-channel dry electrode SSVEP-based biometric approach: Data augmentation techniques against overfitting for RNN-based deep models. *Phys. Eng. Sci. Med.* **45**(4), 1219–1240 (2022).
34. Kamble, S. K. & Sengupta, J. Emotion recognition using wavelet synchrosqueezing transform integrated with ensemble deep learning. *IEEE Sens. J.* https://doi.org/10.1109/JSEN.2023.3335229 (2023).
35. Daubechies, I., Lu, J. & Wu, H. T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* https://doi.org/10.1016/j.acha.2010.08.002 (2011).
36. Bagherzadeh, S. et al. A subject-independent portable emotion recognition system using synchrosqueezing wavelet transform maps of EEG signals and ResNet-18. *Biomed. Signal Process. Control.* https://doi.org/10.1016/j.bspc.2023.105875 (2024).
37. Mumtaz, W. et al. A wavelet-based technique to predict treatment outcome for major depressive disorder. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0171409 (2017).
38. Cavanagh, J. F., Bismark, A. W. & Frank, M. J. Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed EEG. *Comput. Psychiatry* https://doi.org/10.1162/cpsy_a_00024 (2019).
39. Sedehi, J. F., Dabanloo, N. J., Maghooli, K. & Sheikhani, A. Develop an emotion recognition system using jointly connectivity between electroencephalogram and electrocardiogram signals. *Heliyon* https://doi.org/10.1016/j.heliyon.2025.e02855 (2025).
40. Wu, Y., Meng, T., Li, Q., Xi, Y. & Zhang, H. Study on multidimensional emotion recognition fusing dynamic brain network features in EEG signals. *Biomed. Signal Process. Control* **100**(1), 107054. https://doi.org/10.1016/j.bspc.2024.107054 (2025).
41. Fauman, M. A. *Study Guide to DSM-IV-Tr* (American Psychiatric Pub., 2002).
42. Ganapathy, N., Veeranki, Y. R. & Swaminathan, R. Convolutional neural network based emotion classification using electrodermal activity signals and time-frequency features. *Expert Syst. Appl.* https://doi.org/10.1016/j.eswa.2020.113571 (2020).
43. He, K., Zhang, X. & Ren, S. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016). https://doi.org/10.1109/CVPR.2016.90
44. Zheng, X. et al. Ensemble deep learning for automated visual classification using EEG signals. *Pattern Recognit.* https://doi.org/10.1016/j.patcog.2019.107147 (2020).
45. Zhang, Y. et al. Dynamic gesture recognition model based on millimeter-wave radar with resnet-18 and LSTM. *Front. Neurorobot.* https://doi.org/10.3389/fnbot.2022.903197 (2022).
46. Shalbaf, A., Bagherzadeh, S. & Maghsoudi, A. Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals. *Phys. Eng. Sci. Med.* https://doi.org/10.1007/s13246-020-00925-9 (2020).
47. Sokolova, M. & Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* https://doi.org/10.1016/j.ipm.2009.03.002 (2009).
48. Fitzgerald, P. J. & Watson, B. O. Gamma oscillations as a biomarker for major depression: An emerging topic. *Transl. Psychiatry* https://doi.org/10.1038/s41398-018-0239-y (2018).
49. de Aguiar Neto, F. S. & Rosa, J. L. Depression biomarkers using non-invasive EEG: A review. *Neurosci. Biobehav. Rev.* https://doi.org/10.1016/j.neubiorev.2019.07.021 (2019).
50. Nobakhsh, B. et al. An effective brain connectivity technique to predict repetitive transcranial magnetic stimulation outcome for major depressive disorder patients using EEG signals. *Phys. Eng. Sci. Med.* https://doi.org/10.1007/s13246-022-01198-0 (2023).
51. Thibodeau, R., Jorgensen, R. S. & Kim, S. Depression, anxiety, and resting frontal EEG asymmetry: A meta-analytic review. *J. Abnorm. Psychol.* https://doi.org/10.1037/0021-843X.115.4.715 (2006).
52. Shahabi, M. S., Shalbaf, A. & Rostami, R. A convolutional recurrent neural network with attention for response prediction to repetitive transcranial magnetic stimulation in major depressive disorder. *Sci. Rep.* https://doi.org/10.1038/s41598-023-35545-2 (2023).
53. Khan, D. M. et al. Development of wavelet coherence EEG as a biomarker for diagnosis of major depressive disorder. *IEEE Sens. J.* https://doi.org/10.1109/JSEN.2022.3143176 (2022).

Author contributions

S.B.: Conceptualization, Methodology, Software, Investigation, Validation, Writing—original draft, Writing—review & editing. M.R.N.: Conceptualization, Methodology, Investigation. A.G.: Conceptualization, Methodology, Investigation, Writing—original draft. Pouya Tolou Koroshi: Conceptualization, Methodology, Investigation. S.B.H.: Conceptualization, Methodology, Investigation, Software. F.F.: Conceptualization, Investigation, Writing—original draft, Writing—review & editing. A.S.: Conceptualization, Supervision, Writing—review & editing.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval

In this study, two publicly available dataset are used. These first database has been approved by the human ethics committee of the Hospital Universiti Sains Malaysia (HUSM), Kelantan, Malaysia and the second was approved by the University of Arizona. You can download each one through these links: Dataset I: https://figshare.com/articles/dataset/EEG_Data_New/4244171. Dataset II: https://openneuro.org/datasets/ds003478/versions/1.1.0.

Consent to participate

Informed consent was obtained from all individual participants of both databases included in the study.

Additional information

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.