IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 October 2024, accepted 29 October 2024, date of publication 7 November 2024, date of current version 27 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3493140

RESEARCH ARTICLE

SLiTRANet: An EEG-Based Automated Diagnosis Framework for Major Depressive Disorder Monitoring Using a Novel LGCN and Transformer-Based Hybrid Deep Learning Approach

SAGNIK DE[1], (Student Member, IEEE), ANURAG SINGH[2], (Member, IEEE), VIVEK TIWARI[3], (Senior Member, IEEE), HARSHITA PATEL[4], G. N. VIVEKANANDA[4], (Senior Member, IEEE), AND DHARMENDRA SINGH RAJPUT[4]

[1]Institute of Radio Physics and Electronics, University of Calcutta, Kolkata, West Bengal 700009, India
[2]Department of Electronics and Communication Engineering, International Institute of Information Technology Naya Raipur (IIIT Naya Raipur), Naya Raipur, Chhattisgarh 493661, India
[3]Department of Computer Science and Engineering, ABV-Indian Institute of Information Technology and Management Gwalior (ABV-IIITM Gwalior), Gwalior, Madhya Pradesh 474015, India
[4]School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore 632014, India

Corresponding author: G. N. Vivekananda (vivekananda.gn@vit.ac.in)

This work was supported by Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India.

ABSTRACT Major depressive disorder (MDD) is a mental ailment marked by a loss of interest in activities, persistent depression, and hopelessness. MDD has been on the rise in society in recent decades for varied reasons and has spurred suicidal tendencies among individuals. Early detection, continuous monitoring, and effective treatment are crucial for its impact on quality of life and society. EEG signal models the brain's electrical activities and has emerged as a potential tool to assess the depression status of a person. Due to advancements in sensor technology, fast, convenient, and cost-effective EEG acquisition is now possible, resulting in many EEG-based healthcare monitoring applications in recent years. This work proposes an EEG-headset-based smart monitoring system for real-time diagnosis of MDD in the Internet of Medical Things (IoMT) framework. In this study, we proposed a novel Linear Graph Convolution Network-Transformer-based deep learning approach for categorizing MDD through a time-frequency analysis of EEG signals. The Stockwell transform (S-transform) is employed to exploit the spectro-temporal information from the EEG and the resulting 2D representation is then fed into customized Linear Graph Convolution Network for MDD detection. We have utilized the Weighted Focal Binary Hinge Loss function, specifically designed for customized Linear Graph Convolution Network, to improve learning and handle unbalanced input. Subsequently, a novel Transformer model is designed to refine the MDD classification further. The proposed methodology named SLiTRANet, blends spectral analysis with the S-transform, graph-based learning with Linear Graph Convolution Network, and the sequence modeling capability of the Transformer. The proposed SLiTRANet model can be further integrated within an IoMT framework for automated real-time MDD diagnosis using EEG signals. The proposed methodology is evaluated on two publicly available datasets, MODMA and HUSM datasets. The evaluation results demonstrate the superior performance of the proposed SLiTRANet framework against the existing pre-trained and hybrid deep learning models, achieving remarkable accuracy, sensitivity, specificity, and precision rates of 99.92%, 99.90%, 99.95%, and 99.97%, respectively on HUSM dataset followed by an equally good performance on MODMA dataset with an accuracy of 99.68%. The proposed comprehensive approach implemented on two varied datasets highlights significant advancements in depression detection by outperforming state-of-art approaches.

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy.

INDEX TERMS EEG signals, hybrid deep learning network, linear graph convolution network, IoMT, healthcare monitoring, depression classification.

I. INTRODUCTION

Major Depressive Disorder (MDD) is a potent adversary of human well-being that causes significant harm to both people and society. Depression affects approximately 3.8% of the global population, with higher prevalence among adults, especially women, and individuals over 60 years old. Roughly 280 million people worldwide grapple with this condition. Women are 50% more likely to experience depression than men, with over 10% of pregnant or postpartum women affected. Seen as the fourth most common cause of death for individuals aged 15 to 29, suicide tragically claims over 700,000 lives each year [1]. According to these figures, depression has a significant negative influence [25] on people's lives, and more awareness, resources, and therapeutic techniques are urgently needed.

The importance of diagnosing MDD is growing as more research shows a connection between depression and other illnesses. According to recent research, there is a reciprocal association between MDD and long-term conditions such as diabetes, autoimmune disorders, and cardiovascular disease [2]. The MDD patients usually don't discuss their problem within the family or with the doctors mainly due to social stigma. Currently, diagnosing the depression status of a person is a tedious and time-consuming process as psychiatrists take several sessions/days to reach a diagnostic conclusion. Therefore, it is important to develop automated and smart diagnosis tools to facilitate early and faster diagnosis of MDD. With the advancements in sensor technology and computational efficiency in recent years, the Internet of Medical Things (IoMT) and machine learning-based smart healthcare applications have gained momentum. These advanced technologies have the potential to address futuristic healthcare needs such as personalized and remote healthcare, sky-rocketing healthcare costs, doctors scarcity, and lack of medical infrastructure, especially in rural/remote areas. A typical IoMT and deep learning-based framework to diagnose MDD is depicted in Fig. 1. The Electroencephalogram (EEG) headsets facilitate easy and faster signal acquisition, which can be pushed to the cloud through intermediate nodes/computers. The MDD detection task may happen on the cloud, and results may be fetched back to the end-user through a user interface like a smartphone. Diagnosis results may also be fetched by clinicians/hospitals for remote monitoring of the mental status of the patient and provide medical feedback/advice accordingly. Furthermore, developments in neuroimaging methods [5], like functional magnetic resonance imaging (fMRI), are illuminating the neurological foundations of MDD and opening up novel therapeutic and early intervention tools. Given its significance, several research endeavors center on precisely diagnosing MDD in order to tackle its striking impact on both public health and individuals.

EEG serves as a powerful neurophysiological tool for investigating the brain mechanisms underlying MDD [35]. Its high temporal resolution allows researchers to capture even subtle changes in brain wave activity, providing crucial insights into the neurophysiological alterations associated with the disorder. Specifically, EEG can detect variations in spectral power across different frequency bands, which often reflect the emotional dysregulation characteristic of MDD. This ability to monitor rapid brain oscillations offers valuable information on the pathophysiology of the disorder [36]. EEG's portability, affordability, and non-invasive nature [37] make it suitable for large-scale clinical studies and real-world applications, unlike more costly and cumbersome neuroimaging techniques such as fMRI or PET. Moreover, EEG provides a scalable method for identifying MDD by examining brain oscillations, which reflect coordinated neuronal activity spanning both cortical and subcortical regions, including the frontal, temporal, parietal, and occipital lobes. EEG captures abnormalities in MDD by evaluating power amplitudes in specific frequency bands, each linked to distinct cognitive functions. Altered alpha asymmetry in frontal regions, a key indicator in MDD, distinguishes patients from healthy controls [38]. Beta and low gamma (25–80 Hz) powers in fronto-central regions correlate with attention deficits, while intrinsic local beta oscillations in the subgenual cingulate are inversely linked to depressive symptoms [38]. Additionally, in certain contexts, gamma rhythms—neural oscillations ranging from 25 to 140 Hz—have been effective in distinguishing MDD patients from non-depressed controls, with various antidepressant treatments also influencing gamma activity [38]. Additionally, EEG not only contributes to MDD diagnosis but also aids in understanding the brain correlates of other psychiatric conditions similar to MDD like anxiety, psychosis, and cognitive impairments, thus cementing its role in clinical research and practice [43].

In recent years, deep learning has gained significant attention in detection/ classification-related tasks because of its ability to extract automated hidden non-linear features from the input data, which results in higher accuracy. Thus, in the area of EEG based MDD detection, several pioneering innovative deep-learning models have been introduced [39]. DepHNN by Sharma et al. [10] integrates Long Short-Term Memory (LSTM) and convolutional neural network (CNN) architectures. Unlike traditional methods requiring independent feature extraction, DepHNN utilizes LSTM to incorporate long-term dependencies into the CNN framework. Trained on data from 45 patients, DepHNN achieves high accuracy rates of 99.1% with an error rate of 0.2040 in distinguishing depressed individuals from normals. HybridEEGNet introduced by Wan et al. [13] is a novel convolutional neural network featuring two parallel lines. It aims

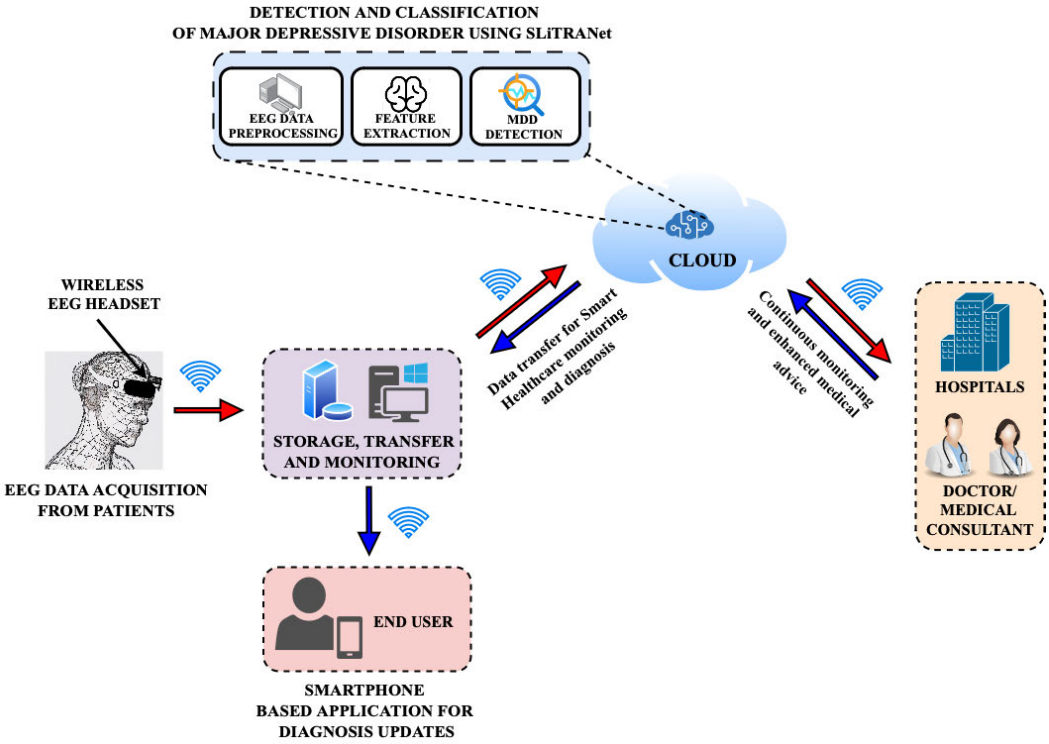

FIGURE 1. An IoMT-deep learning framework for MDD Diagnosis.

to differentiate between medicated and non-medicated MDD patients using synchronous and regional EEG data. Employing ten-fold cross-validation, HybridEEGNet achieves a three-category classification quantified in terms of sensitivity of 68.78%, specificity of 84.45%, and accuracy of 79.08%. Kang et al. [14] propose a method utilizing low-channel EEG signals for early depression diagnosis. Their approach divides signals from four frontal channels into theta, beta, alpha, and delta band frequencies, which are then fed into a convolutional layer-based two-dimensional deep learning model. This method shows promising results with improved sensitivity, specificity, and accuracy. DeprNet introduced by Seal et al. [16] is an 18-layer CNN architecture designed to integrate temporal and spatial features of EEG data. Utilizing Independent Component Analysis (ICA) and filtering techniques to remove artifacts, DeprNet processes 4-second EEG inputs from 19 channels. It achieves notable gains in accuracy compared to previous models, highlighting the differential effects of depression on left and right hemisphere activities. Khan et al. [21] estimate effective connectivity within the brain's default mode network (DMN) using EEG recordings from MDD patients and healthy controls. Their proposed three-dimensional (3D) convolutional neural network demonstrates exceptional performance, achieving 100% accuracy, sensitivity, and specificity in classifying MDD and healthy control participants.

In this work, our suggested model plugs in many significant gaps in most existing works. Although models with excellent accuracy rates, such as HybridEEGNet [13] and DepHNN [10], may not be robust enough to handle imbalances in classes or capture spatial relationships in EEG data. DepHNN may not fully utilize the intrinsic graph structure of EEG signals since it depends on LSTM to incorporate long-term dependencies into CNN. Similarly, HybridEEGNet may miss important spatial correlations across electrode channels in favor of synchronous and local EEG data. Furthermore, while Khan et al.'s 3D convolutional neural network and DeprNet exhibit exceptional 100% performance, they might be overfitted or not dealt with concerns of class imbalance. In contrast, our method uses a Transformer architecture for classification, combines the Stockwell Transform to extract precise time-frequency information, and uses a novel Linear Graph Neural Network (LGCN) with Weighted Focal Binary Hinge (WFBH) loss function to capture spatial dependencies, allowing for improved modeling of long-range dependencies. By integrating these components, our model overcomes the shortcomings of earlier approaches in depression identification and provides improved resilience, accuracy, and interpretability when compared to current techniques.

The major contributions of this work are as follows:

- Introducing an EEG-based automated IoMT framework for smart MDD monitoring using cutting-edge Deep Learning technology.
- Introducing a novel approach for feature extraction in MDD classification by integrating the Stockwell transform for spectro-temporal EEG representation and employing the Linear Graph Neural Network (LGCN)

to analyze EEG data, effectively capturing spatial dependencies within brain networks.
- Introducing the Weighted Focal Binary Hinge (WFBH) Loss tailored for LGCN, effectively addressing imbalanced data and improving model performance in depression classification.
- Incorporating a novel Transformer architecture to refine classification results, leveraging its sequence modeling capability to enhance accuracy and reliability in depression detection.

The structure of this research article unfolds as follows: Section II delves into related works concerning depression classification. In Section III, we provide detailed insights into the dataset and the formulation of our proposed classification architecture. The experimental findings and results of our framework are elucidated in Section IV. Section V serves as the conclusion, summarizing the key findings and highlighting the future scope of the study.

II. RELATED WORKS

A. EXISTING MDD DETECTION APPROACHES

Mumtaz et al. [2] employed resting-state EEG data to evaluate a machine learning (ML) strategy on 33 MDD patients and 30 healthy controls. In order to diagnose depression, the machine learning system used EEG-derived parameters, such as the power of various EEG frequency bands and EEG alpha interhemispheric asymmetry, to distinguish between MDD patients and healthy controls. Mahato and Paul [3] deployed a variety of classifiers, such as quadratic discriminant analysis, radial basis function network (RBFN), linear discriminant analysis (LDA), and multi-layered perceptron neural network (MLPNN). When utilizing linear features with the MLPNN classifier, alpha power has the best classification accuracy of 91.67%. For non-linear features, RWE and WE yielded the highest classification accuracy of 90% using the RBFN and LDA classifiers, respectively. The maximum classification accuracy of 93.33% was attained by combining linear and non-linear characteristics, namely alpha power and RWE, using both MLPNN and RBFN classifiers.

A three-step method for implementing the concept is suggested by Sharma et al. [4] The first step is to extract several important properties from multichannel EEG data, such as coherence, statistical, time-related, linear, fractal dimension, and non-linear features. The second step involves selecting the most relevant characteristics through the use of three feature selection techniques: Neighbourhood Component Analysis (NBA), Relief-based Algorithm (RBA), and Principle Component Analysis (PCA). After that, the performance of several strategies is compared to determine which is the best way to execute the model. The final step involves classifying the participants into normal and depressed groups using five different classifiers. Sharma et al. [5] present DepCap, a state-of-the-art wearable smart cap that provides real-time depression diagnosis via EEG signals. In order to find significant features, the EEG signals of healthy and depressed participants are first converted into spectrogram images using the Short-Time Fourier Transform (STFT). A classification model is then fed these spectrogram images.

Kang et al. [6] suggests a novel method of employing EEG visualization to diagnose depression, emphasizing asymmetry as a critical biomarker. This paper presents a deep-asymmetry methodology that transforms EEG asymmetry information into matrix pictures for input to a CNN, in contrast to conventional methods like STFT and wavelet. This method outperforms many other approaches and achieves an accuracy of 98.85% in the alpha band. Further, Aydemir et al. [7] presented a novel automated approach for diagnosing MDD from EEG waves. The model is divided into three basic stages: creating multileveled features using melamine pattern and discrete wavelet transform (DWT); selecting relevant features using neighborhood component analysis (NCA); and classifying the results using SVM and KNN classifiers. A notable aspect of their research is the application of the melamine pattern, which generates 1536 features using the molecular structure of melamine. Moreover, DWT coefficients are used to derive statistical information.

Saeedi et al. [19], proposed an EEG-centered deep learning architecture in their study to differentiate between people with MDD and healthy people. They begin by extracting the correlations between the EEG channels utilizing sophisticated brain connectivity analysis techniques like Direct Directed Transfer Function (dDTF) and Generalized Partial Directed Coherence (GPDC). For every subject, they generate a distinct visual representation by integrating sixteen connection techniques spread over eight frequency bands. After that, five different deep learning architectures are fed these created EEG signal images to get enhanced performance. This better performance is attributed to the unusual architecture of the model, which is adept at recreating the temporal and spatial correlations observed in brain connection patterns.

B. EXISTING SMART WEARABLE DEVICES FOR HEALTH MONITORING

While there are many devices available for detecting physical disorders in people, there are still few measures available for evaluating an individual's mental state. IoMT-based applications and smart healthcare both depend on the accurate collection and storage of biosignals [29]. In the last ten years, portable consumer electronics have provided fresh insights for researchers studying the Internet of Wearable Devices. For real-time disease monitoring, these devices integrate a variety of biosignals, such as electrocardiography (ECG), electroencephalography (EEG), and photoplethysmography (PPG).

There are several non-intrusive smart wearable gadgets have been devised to assess various aspects of health. An ECG-based wearable device, for example, was suggested

for the purpose of monitoring arrhythmia disorders [30]. Wireless sensor networks have been used to comprehensively study the usage of ECG for monitoring cardiovascular disease [31]. Despite its considerable latency [34], a wearable system for quick seizure detection utilizing EEG was proposed. eSeiz, a different EEG-based consumer electronics product, showed better latency performance [6]. While some EEG features may be useful in detecting depression, there is a dearth of research on the real-time application of smart healthcare equipment for depression diagnosis. A noteworthy study suggested the use of a wearable smart device called Dep-Cap to identify depression [5]. This device uses an algorithm based on deep learning to interpret biosignals.

The review of the literature identifies a number of research gaps in the use of EEG signals for the diagnosis and monitoring of MDD. These gaps involve (a) lack of integration of linear and non-linear features, (b) dependence on manual feature extraction approaches, (c) limited exploration of effective deep learning model architectures, and (d) availability of low latency IoMT devices. The prime objective of this study is to plug in the aforementioned research gaps. First of all, by utilizing the Transformer design to smoothly integrate linear and non-linear features, it improves classification accuracy and permits a more comprehensive comprehension of EEG data. Secondly, it gets beyond the drawbacks of conventional techniques by utilizing the S-Transform, an improved version of STFT and WT that is utilized in signal processing to capture spatio-temporal relationships across EEG channels. To further automate feature extraction and reduce the need for human involvement while increasing efficiency, the model also includes a LGCN. In conclusion, the innovative fusion of the Transformer design with the S-Transform and LGCN offers a distinct and sophisticated method for EEG-based MDD diagnosis required for accurate and reliable monitoring of the mental state of an individual.

III. PROPOSED SMART MDD DETECTION AND MONITORING FRAMEWORK

The improvements in the sensor technology and enhanced on-chip computational power have enabled the IoMT framework to offer numerous e-healthcare opportunities, such as personalized healthcare, and continuous remote monitoring. We proposed a similar framework here for MDD patients for precise and reliable MDD diagnosis and monitoring using the proposed SLiTRANet deep learning architecture. An EEG-headset-based automated MDD monitoring approach is shown in Fig. 2, where EEG signals are captured continuously from the EEG headset and go for feature extraction and depression detection modeling after pre-processing. The computations involved in depression detection may happen locally on a local computer/smartphone taking input EEG wirelessly from the headset or on the cloud by directly pushing the input EEG to the cloud. The MDD diagnosis results may be fetched by the consultant doctor/physician from the hospital/healthcare center for remote monitoring/feedback or by the patient himself/herself using his/her smartphone. The different modules of the proposed framework are discussed in detail below.

A. DATASET AND PREPROCESSING

In the absence of a real-time EEG headset set, the study presented here is evaluated in offline mode on an open-source dataset that was acquired from Hospital Universiti Sains Malaysia (HUSM) [1]. This dataset contains data collected from 64 participants who were carefully chosen from two different groups. There are 40 males and 24 females, ages ranging from 12 to 77. In this dataset, 34 individuals with a mean age of 40.33 years and a standard deviation of 12.861 years were diagnosed with Major Depressive Disorder (MDD). An additional group of thirty healthy controls with an average age of 38.227 years and a standard deviation of 15.64 years were also included. The diagnoses of MDD were conducted in accordance with the Diagnostic and Statistical Manual IV (DSM-IV), with ethical approval secured from the relevant committee, and all participants providing informed consent [9]. EEG data collection adhered to the international 10-20 electrode placement standard, encompassing frontal [12], temporal, parietal, occipital, and central regions with 19 electrodes (Fp1, Fp2, F3, F4, F7, F8, Fz, T3, T4, T5, T6, P3, P4, Pz, O1, O2, C3, C4, and Cz). For data preprocessing, first, a Butterworth bandpass filter (0.5 Hz - 70 Hz) was applied to the EEG signals to retain relevant neural activity while removing low-frequency drifts and high-frequency noise. To eliminate power line interference, a 50 Hz notch filter was used. Following the filtering, Independent Component Analysis (ICA) [44] was employed to address common artifacts, [18] such as eye blinks, muscle movements, and patient motion. ICA decomposes the multichannel EEG data into statistically independent components, each representing a different signal source. The independent components were then inspected, and those corresponding to artifacts (e.g., large slow-wave components for eye blinks or sharp components for muscle activity) were identified and removed. Finally, the cleaned components were recombined to reconstruct the artifact-free EEG signals. This process ensures that the data used for further analysis is free from non-neural artifacts, thereby improving the accuracy of feature extraction and classification. After preprocessing step, each 5-minute EEG recording was segmented into 10-second intervals, with Z-score normalization employed for amplitude scaling before input to the neural network. Further segmentation divided the 5-minute dataset into epochs of 4 seconds (1024 samples) each, ensuring consistent labeling for subsequent machine learning applications.

Additionally, we have also used another EEG-based depression dataset that has restricted public availability, called MODMA dataset [26], to validate our proposed depression detection framework (S. Transform+LGCN-Transformer). In this database, there were 52 participants:

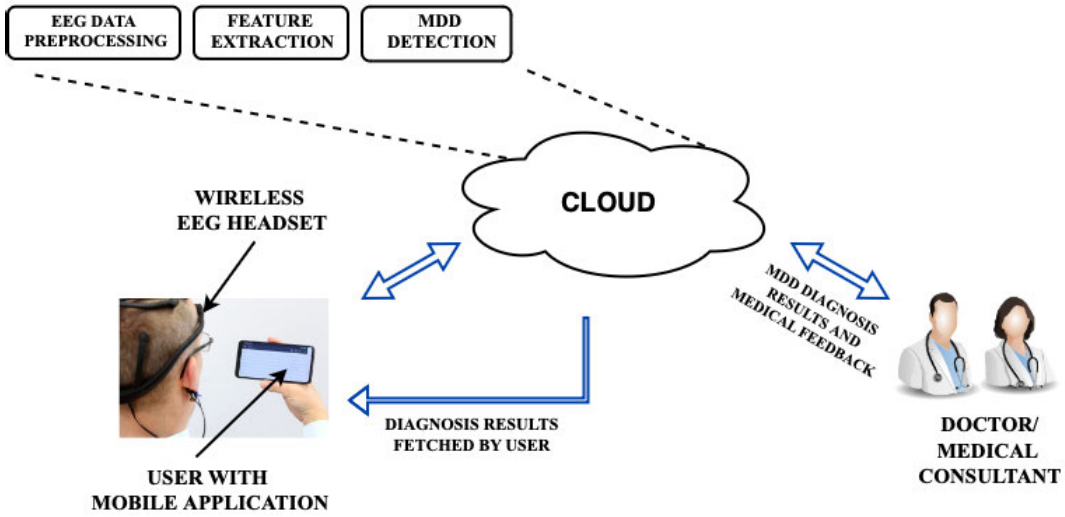

FIGURE 2. The proposed IoMT-based MDD diagnosis and monitoring framework incorporating the SLiTRANet deep learning architecture.

29 healthy controls (20 males and 9 females, ages 18–55) and 23 individuals with depressive diagnoses (16 males and 7 females, ages 16–56). For our study, we initially explored the full-brain 128-electrode EEG data mode. Nevertheless, we restricted our focus to 19 distinct electrodes (Fp1/2, F3/4, F7/8, Fz, C3/4, Cz, P3/4, Pz, O1/2, T3/4, T5/6) in order to optimize time efficiency and computational resources. These electrodes have been widely utilized in the study of MDD, thus this deliberate choice was not only rational but also consistent with other studies [41], [42], which makes them quite pertinent to our work. The aforementioned preprocessing steps (normalization, removal of artifacts using ICA, and segmentation) have also been applied to this database records to prepare them for further processing. A schematic block diagram of the proposed MDD detection approach is shown in Fig. 3. It can be observed that there are three major modules in the approach: preprocessing, feature extraction, and deep learning modeling. Each of the modules is discussed in detail in subsequent sections.

B. STOCKWELL TRANSFORM-BASED T-F ANALYSIS OF EEG SIGNAL

The Stockwell Transformation, or S-transform, proposed by Stockwell et al. in 1996, amalgamates the advantages of Short Time Fourier Transform (STFT) [32] and Wavelet Transform (WT), garnering significant interest across various scientific and engineering domains such as biomedical imaging and signal processing. Unlike traditional methods, the S-transform offers frequency-dependent resolution while maintaining a direct correlation with the Fourier spectrum. Essentially, it serves as a phase-corrected version of the WT, furnishing more accurate insights into a signal's local features during time-frequency (T-F) analysis. This property renders the S-transform particularly robust against non-stationary signals and often provides superior time-frequency resolution [40]. This is why S-transform is chosen for T-F analysis of EEG signal for MDD detection in this work.

Given $\rho \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\int_{-\infty}^{\infty} \rho(t)dt = 1$, the S-transform of a signal $y(t)$ in $L^2(\mathbb{R})$ with respect to the window function $\rho(t)$ is defined by equation (1):

$$S_y(t,f) = \int_{-\infty}^{\infty} y(\tau)\rho^*(\tau - t)e^{-2\pi if\tau}d\tau \tag{1}$$

where,
$S_y(t,f)$ represents the S-transform of $y(t)$ at time t and frequency f.
$\rho^*(t)$ denotes the complex conjugate of $\rho(t)$.
f denotes frequency while t represents time.

The inverse Stockwell Transform, reconstructing the original signal, is given by equation (2):

$$y(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_y(\tau,f)\rho(t-\tau)\exp(2\pi ift)d\tau df \tag{2}$$

This equation represents the reconstruction of the original signal $y(t)$ by integrating over the time-frequency domain, with $g(t-\tau)$ representing a suitable reconstruction window function.

At zero frequency $f = 0$, the Stockwell Transform equals the signal's average as shown in equation (3):

$$S_y(t,0) = \int_{-\infty}^{\infty} y(\tau)d\tau \tag{3}$$

This signifies that at zero frequency, the S-transform effectively computes the average value of the signal over time, providing insight into its overall magnitude or energy distribution.

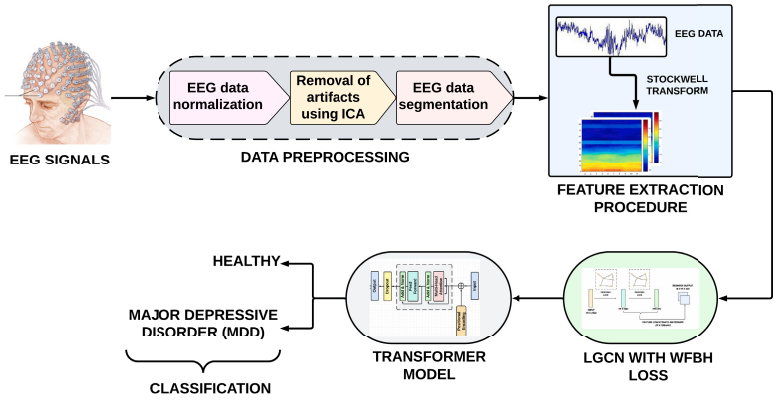

FIGURE 3. Block-diagram of the proposed pipeline of the MDD diagnosis using SLiTRANet.

The Stockwell Transform can also be formulated in the frequency domain as depicted in equation (4):

$$S_y(t,f) = \int_{-\infty}^{\infty} Y(\alpha + f)\hat{\rho}(\alpha/|f|)e^{j2\pi\alpha t}d\alpha \tag{4}$$

Here, $Y(\alpha + f)$ represents the Fourier spectrum of the signal shifted by the frequency f, and $\hat{\rho}(f)$ denotes the frequency domain representation of the window function $\rho(t)$. This formula shows how the S-transform converts a signal from the time domain to the frequency domain, allowing for an in-depth analysis of the signal's frequency content over time.

C. LINEAR GRAPH CONVOLUTION NETWORK (LGCN)

This section elaborates on the structure of the proposed LGCN architecture used for depression-relevant feature extraction.

1) GRAPH CONVOLUTION NETWORK IN SPECTRAL DOMAIN

The complex network of the brain influences diverse mental states, including depression, through connections between distinct brain regions. Understanding the patterns of brain activity linked to depression can be gained by analyzing EEG signals obtained from electrodes applied to the scalp. Through Pearson correlation analysis, we construct a graph where EEG signal features represent nodes, reflecting raw EEG data. The adjacency matrix of this graph, derived from correlation analysis and thresholding, signifies connectivity between nodes. Inspired by Zhao et al.'s [23] research, we propose a depression detection model employing Linear Graph Convolutional Networks (LGCNs). The flow diagram of a generalized LGCN is as depicted in Fig. 4

LGCNs, similar to Graph Convolutional Networks (GCNs), iteratively process node neighborhoods. Utilizing two LGCN layers, each node aggregates features from its neighboring nodes, updating its representation. This aggregation process, which combines node features with neighboring information, is described by the convolutional operation in equation (5):

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{5}$$

Here, $\tilde{A}$ represents the adjacency matrix with added self-loops, $\tilde{D}$ is the degree matrix, $W^{(l)}$ denotes trainable weights, and $\sigma(\cdot)$ applies an activation function, commonly ReLU.

Exploring the relationships between EEG electrode pairs in the International 10-20 system, our LGCN model operates in the Fourier domain. Utilizing convolution kernels g_θ and raw EEG signals x, we perform spectral convolutions as shown in equation (6):

$$g_\theta \star x = U g_\theta U^\top x \tag{6}$$

Here, U signifies the eigenvector matrix derived from the symmetric normalized Laplacian operator, which is computationally intensive. To address this, we employ k-order Chebyshev polynomials [23], approximating convolutions as:

$$g_{\theta'} \star x \approx \sum_{k=0}^{K} \theta'_k T_k(\tilde{L})x \tag{7}$$

where $\tilde{L}$ is a rescaled Laplacian matrix. This approximation significantly reduces computational complexity to $\mathcal{O}(|\mathcal{E}|)$.

Incorporating Chebyshev's linear model, spectral convolutions simplify to:

$$g_{\theta'} \star x \approx \theta\left(I_N + D^{-\frac{1}{2}}AD^{-\frac{1}{2}}\right)x \tag{8}$$

However, direct application within deep neural networks may result in gradient instability. Hence, we normalize the operation by replacing $(I_N + D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$ with $\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$, yielding a layer-wise linear form GCN model expressed as:

$$H^{(l+1)} = f\left(H^l, A\right) = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right) \tag{9}$$

2) WEIGHTED FOCAL BINARY HINGE (WFBH) LOSS FUNCTION

In depression classification, handling imbalanced data is crucial to ensure the model effectively learns from both

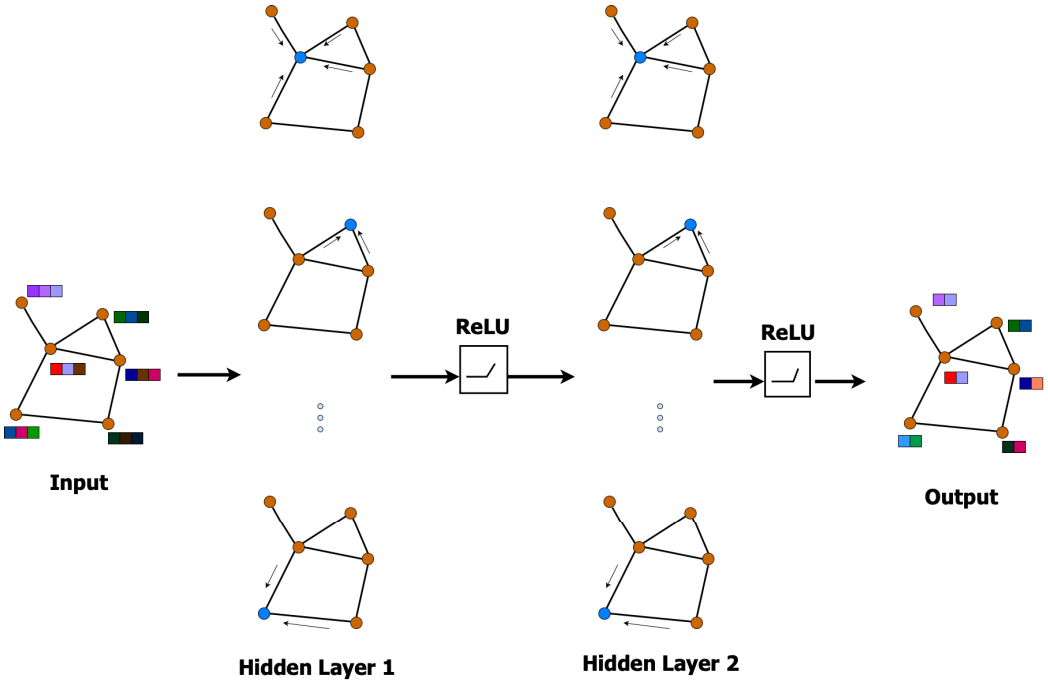

FIGURE 4. The flow diagram of a generalized Linear Graph Convolution Network.

positive and negative samples. The original loss function used in Linear Graph Convolutional Networks (LGCNs) is cross-entropy loss [23], given by:

$$\mathrm{CE}_L(z, y) = \begin{cases} -\log(z) & \text{if } y = 1 \\ -\log(1 - z) & \text{otherwise} \end{cases} \tag{10}$$

Here, z represents the predicted probability of a class and y is the true label. However, in scenarios with imbalanced data, this loss function may overly prioritize the majority class, leading to suboptimal performance.

To address this issue, focal loss, initially proposed for object detection tasks, can be employed. Focal loss(F_L) is formulated as:

$$\mathrm{F}_L(z) = -\alpha_z(1 - z)^{\gamma} \log(z) \tag{11}$$

In this formulation, z represents the predicted probability, and α_z adjusts the weight of positive and negative samples. The parameter γ allows for controlling the emphasis on challenging samples.

Weighted Focal Binary Hinge (WFBH) Loss can be applied to improve model performance even more in imbalanced settings. Combining binary hinge loss and focal loss, WFBH focuses on difficult samples and provides adaptability against class imbalance. Its formulation is given by equation (12):

$$\mathrm{WFBH}(z, y) = \beta \cdot \mathrm{H}_L(z, y) + (1 - \beta) \cdot \mathrm{F}_L(z, \gamma) \tag{12}$$

The binary hinge loss in this case is represented by $\mathrm{H}_L(z, y) = \max(0, 1 - (2y - 1) \cdot z)$ and the focal loss by $\mathrm{F}_L(z, \gamma)$. Due to the parameter β, which balances the contributions of hinge and focal loss, WFBH is appropriate for tasks involving the categorization of depression, where difficult samples and imbalanced data are frequently encountered.

The working mechanism of the proposed LGCN architecture is as described below. Each node in the first layer enhances its own understanding by combining the attributes of the nodes that are close to it. Each node's features in the second layer are a combination of its own characteristics and those of its neighbors. Each node continues this process, honing its features in response to its environment. They can then be utilized to the Transformer model for classification by merging the characteristics from both layers and reshaping the output of LGCN module. The overview of the novel LGCN with WFBH is as illustrated in Fig. 5. Furthermore, the algorithm of the proposed LGCN with WFBH is as given in Algorithm 1.

D. PROPOSED NOVEL TRANSFORMER ARCHITECTURE

The proposed novel Transformer model for classifying EEG signals into healthy and MDD classes begins by embedding the 2D input matrix representing the EEG signal. This embedding process transforms the input matrix X into a sequence of embeddings E, ensuring compatibility with the subsequent layers. The embedding operation can be mathematically expressed as:

$$E = \mathrm{Embed}(X) \tag{13}$$

Next, the self-attention mechanism [24] is employed to compute attention scores between each pair of time

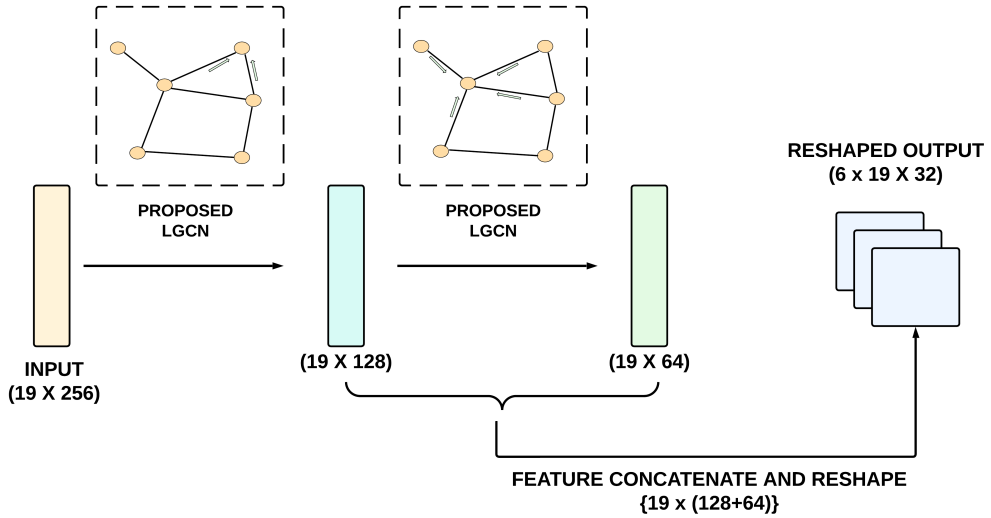

FIGURE 5. The block diagram of the proposed LGCN architecture with WFBH loss function.

Algorithm 1 Algorithm for Proposed Novel LGCN With WFBH Loss

Input: EEG data represented as graph $G = (V, E)$, with features X and adjacency matrix A

Output: Predicted labels for depression classification

1: **Compute** the degree matrix $\tilde{D}$ from $\tilde{A}$
2: **Initialize** parameters: trainable weights $W^{(l)}$ for each layer l
3: **Compute** the rescaled Laplacian matrix $\tilde{L}$ from the adjacency matrix A
4: **for** each layer l in LGCN **do**
5: **Compute** spectral convolutional operation:
6: $H^{(l+1)} = \sigma\left(\theta\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right)\right)$
7: **Update** node representations using activation function $\sigma(\cdot)$
8: **end for**
9: **Compute** Weighted Focal Binary Hinge (WFBH) Loss:
10: $\mathrm{F}_L(z) = -\alpha_z(1-z)^\gamma \log(z)$
11: $\mathrm{H}_L(z, y) = \max(0, 1 - (2y-1)\cdot z)$
12: $\mathrm{WFBH}(z, y) = \beta \cdot \mathrm{H}_L(z, y) + (1-\beta)\cdot \mathrm{F}_L(z, \gamma)$
13: **Update** parameters using gradient descent with WFBH Loss
14: **return** Predicted labels

steps in the EEG signal. This enables the model to weigh each pertinent time step's significance in the classification procedure. The attention score calculation is represented by:

$$A_{ij} = \text{softmax}\left(\frac{Q_i K_j^T}{\sqrt{d_k}}\right)V \tag{14}$$

Multi-head attention (MHA) is then utilized to capture diverse patterns and dependencies within the EEG signal. By running the self-attention mechanism multiple times in parallel with different linear projections, the model can extract rich representations from the input data. The multi-head attention [24] operation is mathematically expressed as:

$$\begin{aligned}\text{MHA}\,(Q, K, V) &= \text{Concat}\,(\text{head}_1, \ldots,\ \text{head}_h)\,W^O \\ \text{where head}_i &= \text{Attention}\left(QW_i^Q, K\ W_i^K, V\ W_i^V\right)\end{aligned} \tag{15}$$

The output of the MHA layer is passed through a position-wise feedforward network (FFN) [24] to capture complex interactions within the EEG signal. A novel aspect of this model is the computation of L1 norms (Manhattan norms) for the probabilistic outputs from all attention heads. The L1 norm for each attention head is computed as:

$$\text{L1 Norm}(A_i) = \sum_j |A_{ij}| \tag{16}$$

The attention head with the highest L1 norm is considered to have the highest attention probability. This specific attention output is selected as the most relevant and used as a residual connection to refine the final projected attention output (F_{AO}):

$$\mathrm{F}_{AO} = \text{Attention Output} + \text{Attention Output}_{\text{highest L1 Norm}} \tag{17}$$

Incorporating this highest norm attention output as a residual connection ensures that the most pertinent features are emphasized, thereby enhancing the overall feature representation and classification accuracy.

This network applies two linear transformations followed by a ReLU activation function. The FFN operation can be denoted as:

$$\text{FFN}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2 \tag{18}$$

Residual connections and layer normalization [24] are then applied to facilitate smooth gradient flow and alleviate

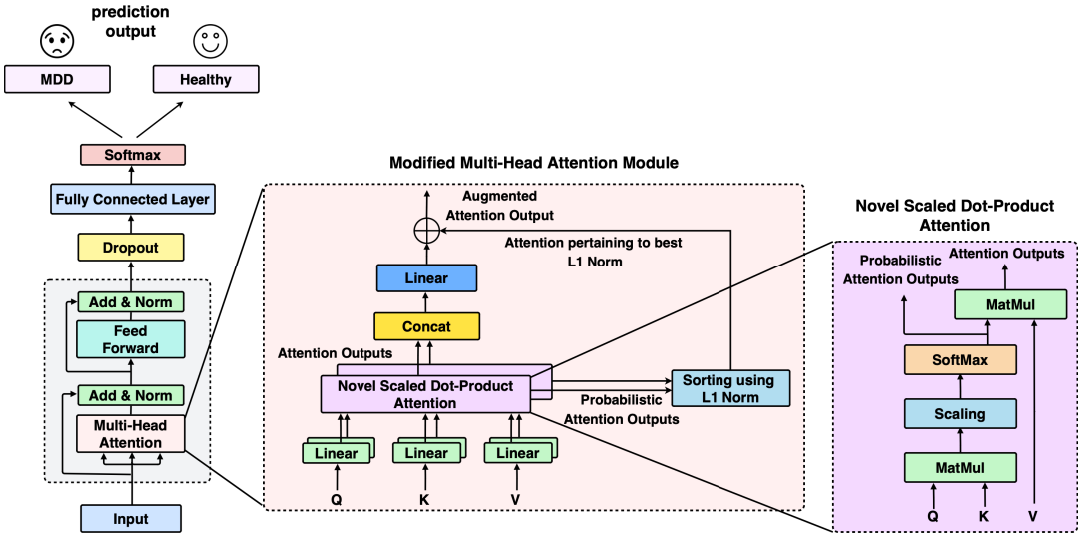

FIGURE 6. The schematic of the proposed novel Transformer model for MDD diagnosis.

vanishing gradient issues during training. These connections ensure that information from earlier layers is preserved and propagated effectively through the network. The layer normalization operation is represented as:

$$\text{LayerNorm}(x + \text{Sublayer}(x)) \tag{19}$$

Finally, the output of the Transformer is passed through a fully connected layer with softmax activation for binary classification into healthy and MDD classes. This layer produces the final classification probabilities based on the learned representations from the input EEG signal. A schematic diagram of the proposed Transformer model is illustrated in Fig. 6

E. PROPOSED DEEP LEARNING ARCHITECTURE FOR MDD DIAGNOSIS AND MONITORING

The input EEG signal is first represented as a 19 $\times$ 2560 matrix, including 10 seconds (after data segmentation) of data sampled at a rate of 256 samples per second over 19 electrode channels as per HUSM dataset. The signal is transformed after applying the S-transform, which splits the time-frequency plane into 256 time bins and 19 frequency bins. The original matrix is compressed into a 256 $\times$ 19 matrix using this transformation, with each member denoting a distinct time-frequency bin. The resulting matrix effectively captures the EEG signal's time-varying spectral properties. As a result, a smaller 19 $\times$ 256 matrix is used as the LGCN's input. The LGCN comprises two layers, generating 128 and 64 frequency maps, respectively. These feature maps are concatenated and reshaped, resulting in an output matrix with a total of 192 feature maps and dimensions of 6 $\times$ 19 x 32. This matrix contains spatial information. The Transformer receives this output, which is a reflection of the processed features that were taken out of the EEG data. Together with Multi-Head Attention (MHA) modules, self-attention mechanisms within the Transformer enable the learning of representations over several channels and time steps. To further improve learning, the encoder also includes Add and Norm layers. These representations are iteratively improved by the Transformer across several layers, culminating in dense layers for additional processing. A dropout operation with a dropout rate of 0.5 is also applied to prevent overfitting. The output dimension is flattened and passed to a dense layer with softmax activation for final binary classification (i.e. Healthy or MDD). Table 1 gives the optimal hyperparameters of the LGCN and Transformer architectures used in this work.

The step-by-step procedure of the proposed SLiTRANet for MDD detection is given in Algorithm 2.

IV. EXPERIMENTAL RESULTS

A. HARDWARE SPECIFICATIONS AND ENVIRONMENTAL SETUP

Using Python 3.8.1 and Jupyter Notebook and the Keras framework, deep learning model training and evaluation were done. The system used for implementation has following configurations: an Intel Core i7 processor, an NVIDIA GTX 1650 GPU that supports CUDA, and 16GB of RAM to enable effective model training and assessment. This setting is used to customize and improve the model architecture throughout the course of a 50-epoch training regimen. Each epoch begins with a randomization of the dataset to ensure a varied exposure to the data and support robust learning. Using an initial learning rate of 0.0001, the ADAM (Adaptive Moment Estimation) optimizer [33] is used to orchestrate the model's optimization with a batch size of 128.

B. TRAINING, VALIDATION AND TESTING PROCEDURE

The development of robust classifier models in machine learning depends on the efficient preparation of training, validation, and testing datasets. Validation data helps

TABLE 1. Optimal hyperparameters for LGCN and transformer.

Hyperparameters	LGCN	Transformer
Input dimension	19 x 256	6 x 19 x 32
Output dimension	6 x 19 x 32	2
Number of feature maps	192	-
Number of time steps	256	-
Number of features	19	-
Kernel size	5x5	-
Number of layers	2	6
Activation function	ReLU	ReLU
Pooling	Max pooling (2x2)	-
Number of attention heads	-	8
Hidden dimension	-	512
Dropout rate	-	0.5

Algorithm 2 Proposed SLiTRANet Algorithm

Input: Preprocessed EEG data represented as X

Output: Predicted labels for Healthy and MDD classes

1: **Apply** Stockwell Transform to X to obtain time-frequency representation S_X
2: **Feed** S_X into Linear Graph Convolution Network (LGCN) with Weighted Focal Binary Hinge (WFBH) Loss:
3: $LGCN_{\text{output}} \leftarrow \text{LGCN}(W_{\text{LGCN}}, S_X)$
4: **Reshape** $LGCN_{\text{output}}$ into a suitable format:
5: $LGCN_{\text{reshaped}} \leftarrow \text{Reshape}(LGCN_{\text{output}})$
6: **Feed** $LGCN_{\text{reshaped}}$ into Transformer for classification into Healthy and MDD classes:
7: $PredictedLabels \leftarrow \text{Transformer}(W_{\text{Transformer}}, LGCN_{\text{reshaped}})$
8: **return** $PredictedLabels$

minimize overfitting by adjusting model hyperparameters, whereas training data serves as the basis for identifying patterns and generating predictions. Evaluation of the model's performance on omitted data is done using testing data. Random data splitting and subject-based data splitting are the two primary methods for preparing datasets. Given the limited subject representation in topic-based splitting, bias and variability may be introduced, whereas representative samples and various characteristics are guaranteed by random splitting. 70% of the dataset samples are used for training, 10% for validation, and 20% for testing using a stratified random splitting technique, which guarantees variation in the SLiTRANet framework's training process. This strategy improves model performance and prevents overfitting, when combined with a 10-fold cross-validation method. Table 3 and Table 4 demonstrates that allocating 70% of the dataset for training, 10% for validation, and 20% for testing yields the highest accuracy for both datasets. Consequently, this split ratio was selected for training, validation, and testing.

C. PERFORMANCE EVALUATION METRICS

We employed a wide range of performance indicators in our trials to evaluate the effectiveness of our suggested framework. These metrics include accuracy [11], specificity [20], precision [17], and sensitivity [22] (recall), which, taken together, provide a more comprehensive assessment of our model's effectiveness. Equations (20) through (25) provide the mathematical formulations for the above performance metrics.

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{20}$$

$$\text{Specificity} = \frac{T_N}{T_N + F_P} \tag{21}$$

$$\text{Precision} = \frac{T_P}{T_P + F_P} \tag{22}$$

$$\text{Sensitivity (Recall)} = \frac{T_P}{T_P + F_N} \tag{23}$$

These performance metrics collectively offer a comprehensive assessment of our model's effectiveness in depression classification, providing valuable insights into its capabilities and limitations. We have performed a 10-fold cross-validation during model evaluation and reported the averaged cross-validation results.

D. PERFORMANCE ASSESSMENT OF PROPOSED MDD DETECTION FRAMEWORK

This section presents the performance evaluation of the proposed approach for the two classes, Major Depressive Disorder (MDD) and Healthy in terms of different metrics such as specificity, accuracy, sensitivity, and precision. First, the proposed deep learning model was trained properly. On both datasets, setting the learning rate to 0.001 initially produced high accuracy; however, the validation set revealed significant variations in accuracy and loss, suggesting inadequate training. Training accuracy increased to 0.9992 and 0.9968 for the HUSM and MODMA datasets, respectively, after the learning rate decreased to 0.0001. The training loss values for the HUSM and MODMA datasets were approximately 0.09 and 0.36, respectively. This change in the learning rate along with the Adam optimizer, and a batch size of 128, produced optimal results and stabilized the model's training/validation performance on both datasets. The algorithm was run for 50 epochs. The training and validation accuracy curves, as well as the training and validation loss curves versus the number of epochs, are

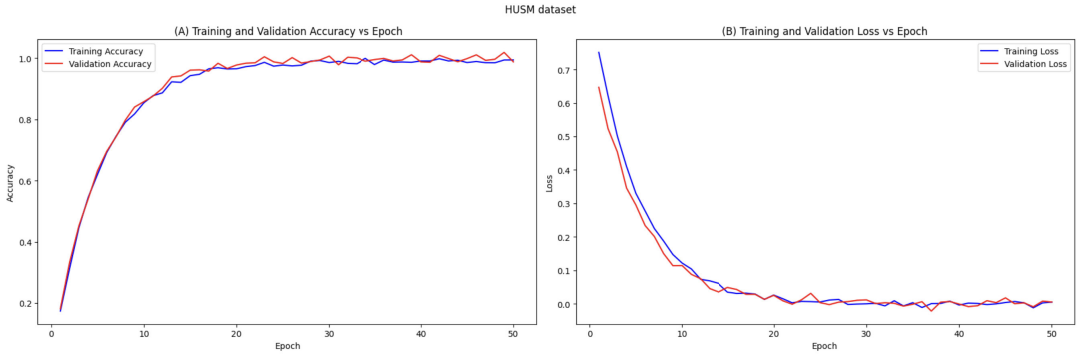

FIGURE 7. The (A) Training and validation accuracy and (B) Training and validation loss curves using HUSM dataset.

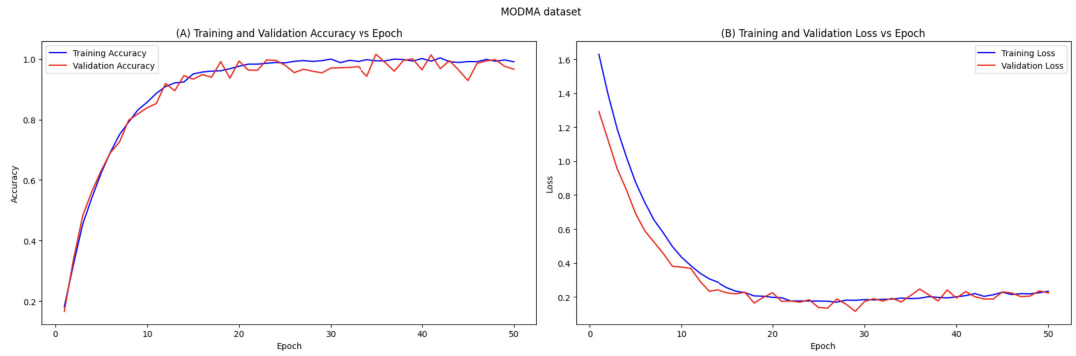

FIGURE 8. The (A) Training and validation accuracy and (B) Training and validation loss curves using MODMA dataset.

depicted in Fig. 7 for the HUSM dataset and in Fig. 8 for the MODMA dataset.

Table 2 depicts the values of different performance metrics obtained for the two evaluated databases. The maximum accuracy was obtained for MDD class (99.96%) followed by healthy class (99.88%). A similar trend was obtained for the MODMA dataset also, where the maximum accuracy of the MDD class was found to be 99.76%. The average performance metrics for both classes are found to have an accuracy of 99.92%, sensitivity of 99.90%, specificity of 99.95%, and precision of 99.97%. These statistics demonstrate the robust performance of the proposed detection model in accurately distinguishing between Healthy and MDD instances. The confusion matrix [27] of the proposed framework on both HUSM and MODMA dataset is also computed and illustrated in Fig. 9. Since the proposed framework performs best on the HUSM dataset, all the experimentations, here onwards, have been performed taking HUSM dataset only.

E. PERFORMANCE ANALYSIS OF PROPOSED DEPRESSION DETECTION FRAMEWORK ACROSS DIFFERENT DATA SPLITTING EXPERIMENTS

Further, to find out the best data splitting ratio for training, testing, and validation, we have also assessed the performance of the proposed SLiTRANet model across different data-splitting experiments. Table 3 and Table 4 display the performance of the HUSM and MODMA datasets when trained, validated, and tested using various splitting approaches. Our findings indicate that our model demonstrates superior performance on the MODMA dataset, thereby establishing the effectiveness of the proposed framework over other existing models. As shown in both the tables, the best performance is achieved with the ratio 70:10:20. Therefore, in this study, the designed framework is trained, validated, and tested using randomly selected 70%, 10%, and 20%, respectively. It can also be seen that subject-based splitting produces comparable results on both datasets, demonstrating the model's strong generalization ability.

F. CROSS-DATASET VALIDATION OF THE PROPOSED MODEL

To demonstrate the robustness of the proposed model on unseen data, we have also carried out cross-dataset validation, where training was done on one dataset (HUSM) and testing was done on another unseen dataset (MODMA) and vice-versa. This is a very strict validation protocol but very important to highlight the model's generalizability and model's capability to handle inter-patient variability, especially in EEG-based applications [45], [46]. Table 5 presents the performance of classifiers on different datasets

TABLE 2. Performance assessment of the proposed depression detection framework on two datasets.

Class	HUSM Dataset Metrics (%)				MODMA Dataset Metrics (%)			
	Accuracy	Sensitivity	Specificity	Precision	Accuracy	Sensitivity	Specificity	Precision
Healthy	99.88	99.95	99.82	99.97	99.60	99.55	99.70	99.67
MDD	99.96	99.85	99.88	99.97	99.76	99.75	99.68	99.83
Average	**99.92**	**99.90**	**99.95**	**99.97**	**99.68**	**99.65**	**99.70**	**99.75**

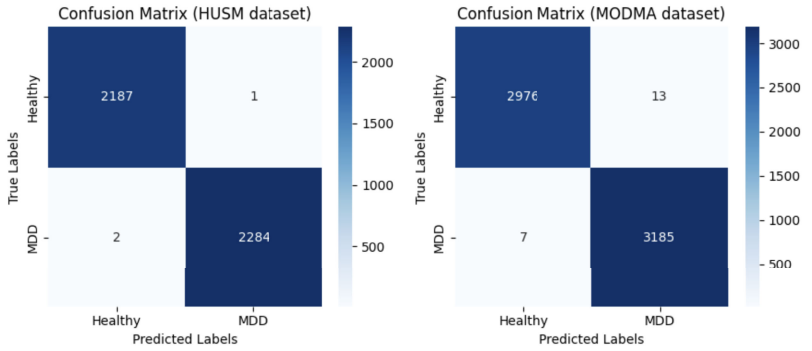

FIGURE 9. The confusion matrix of the proposed Depression Detection framework on testing using both datasets.

TABLE 3. Performance of the proposed model on testing data while trained, validated, and tested using different data splitting approaches (HUSM dataset).

Data splitting	Data splitting ratio	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
Random	**70: 10: 20**	**99.92**	**99.97**	**99.90**	**99.95**
Random	60: 10: 30	99.25	99.41	99.33	99.28
Random	40: 10: 50	94.75	94.34	93.74	94.55
Random	50: 10: 40	95.58	96.33	95.97	96.39
Random	55: 20: 25	96.19	96.56	96.71	96.33
Subject-based	42: 6: 12	98.88	98.76	98.82	98.87
Subject-based	36: 10: 14	92.74	92.33	92.64	92.73
Subject-based	28: 14: 18	91.56	91.66	91.45	91.33
Subject-based	24: 16: 20	90.33	90.75	90.54	91.16

TABLE 4. Performance of the proposed model on testing data while trained, validated, and tested using different data splitting approaches (MODMA dataset).

Data splitting	Data splitting ratio	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
Random	**70: 10: 20**	**99.68**	**99.65**	**99.70**	**99.75**
Random	60: 10: 30	98.66	98.72	98.25	98.33
Random	40: 10: 50	93.75	94.24	93.84	93.66
Random	50: 10: 40	96.86	96.55	96.66	96.68
Random	55: 20: 25	95.66	95.75	95.28	95.35
Subject-based	39: 6: 8	96.48	96.66	96.33	96.75
Subject-based	32: 9: 12	91.74	92.13	91.64	91.33
Subject-based	26: 12: 15	90.41	90.66	90.33	90.25
Subject-based	21: 14: 18	89.63	90.86	89.54	89.33

through cross-dataset validation (CD_{val}), training on 90% of one database and testing on 10% of another. This approach is essential as it ensures the model's generalizability and robustness across diverse datasets, preventing overfitting to a single dataset. In Experiment CD_{val1}, the model trained on the HUSM dataset and tested on MODMA achieved an average accuracy of 97.10%, demonstrating strong performance in detecting both Healthy and MDD classes. Conversely, in Experiment CD_{val2}, where the training and testing datasets were swapped, the model obtained an average accuracy of 96.02%. The model was adapted to each dataset to ensure uniformity during analysis, confirming that while performance may vary slightly due to dataset characteristics, the model is capable of generalizing effectively across different datasets.

TABLE 5. Performance overview of classifiers on different datasets.

Experiment	Train Dataset	Test Dataset	Class	Accuracy (%)
CD_{val1}	HUSM	MODMA	Healthy	96.72
			MDD	97.48
			Average	**97.10**
CD_{val2}	MODMA	HUSM	Healthy	95.83
			MDD	96.21
			Average	**96.02**

G. CROSS-SUBJECT VALIDATION OF THE PROPOSED MODEL

Cross Subject Validation (CS_{val}) is a technique used to assess the robustness and generalization ability of a machine

TABLE 6. Performance overview of proposed architecture on cross-subject validation.

Experiment	Dataset Used	Class	Average Accuracy (%)
CS_{val1}	HUSM	Healthy	97.68
		MDD	98.42
		Average	**98.05**
CS_{val2}	MODMA	Healthy	96.34
		MDD	97.82
		Average	**97.08**

TABLE 7. Performance overview of LGCN(WBFH)-Transformer against other LGCN variants.

Model	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)
LGCN-Transformer	94.98	95.20	94.90	94.30
LGCN(Focal Loss)-Transformer	97.32	96.80	97.93	98.50
LGCN(WFBH)-Transformer	**99.92**	**99.97**	**99.95**	**99.90**

learning model in handling variations across different individuals (subjects). In CS_{val}, the dataset is divided such that one subject's data is held out for testing, while the data from the remaining subjects is used for training the model. This process is repeated for every subject, and the test results are averaged to obtain the final accuracy. CS_{val} is particularly important in real-world applications because human data often exhibits high inter-subject variability, making subject-independent testing crucial for model reliability. The importance of CS_{val} lies in its ability to simulate real-world conditions where a model trained on a specific group of individuals must generalize well to unseen subjects. This validation method ensures that the model does not overfit to individual-specific patterns but captures generalized features relevant to all subjects. In Table 6, two experiments were performed: CS_{val1} on the HUSM dataset and CS_{val2} on the MODMA dataset. For CS_{val1}, the model achieved an average accuracy of 98.05%, with 97.68% accuracy for classifying healthy subjects and 98.42% for MDD subjects. Similarly, CS_{val2} showed 97.08% average accuracy with healthy subjects at 96.34% and MDD subjects at 97.82%, indicating the model's strong performance across datasets.

H. IMPORTANCE OF THE PROPOSED LGCN CUSTOMIZATION IN IMPROVING THE DETECTION ACCURACY

To highlight the importance of the proposed customizations in the base model of LGCN-Transformer, we have compared the base model and the focal loss model of the deep learning architecture in this section. The comparison of different versions of the LGCN-Transformer for detecting MDD offers valuable insights and emphasizes the role of customization in achieving better performance. Table 7 illustrates the performance of the aforementioned LGCN variants with the Transformer. At first, the LGCN-Transformer base model shows good performance, with well-balanced precision, specificity, and sensitivity, and achieves an accuracy of 94.98%, suggesting its dependability. When Focal Loss is included, performance improves. This is demonstrated by the LGCN(Focal Loss)-Transformer in Table 7, with an accuracy of 97.32% and improved sensitivity and precision, which are essential for precise detection tasks. When the loss function changes to Weighted Focal Binary Hinge (WFBH) loss as done in the proposed architecture, presented as LGCN(WFBH)-Transformer, MDD detection performance improved significantly with high precision, specificity, sensitivity, and accuracy (99.92%) as depicted in the comparison table, outperforming earlier models by substantial margins. These results highlight LGCN-based models' utility and emphasize the importance of incorporating WFBH loss function for additional performance improvement.

I. ABALATION ANALYSIS

The ablation analysis (performed on HUSM dataset) presented in Table 8 highlights the impact of removing or modifying key components of the proposed SLiTRANet model on its performance metrics, including accuracy, precision, specificity, and sensitivity. The SLiTRANet model achieves an exceptional accuracy of 99.92%, precision of 99.97%, specificity of 99.95%, and sensitivity of 99.90%, indicating its robust MDD classification capabilities. The removal of the S-Transform, which transforms raw data inputs into 2D T-F representation, leads to a substantial accuracy drop to 90.28%, underscoring the S-Transform's vital role in effectively extracting MDD-relevant features. Replacing the LGCN with a conventional CNN results in further performance degradation, resulting in an accuracy of 84.72%, emphasizing the LGCN's ability to model complex spatial relationships. Notably, the elimination of the novel Transformer reduces accuracy to 85.72%, revealing that its unique architecture enhances contextual feature extraction beyond that of a baseline Transformer, which achieves only 89.38% accuracy. Moreover, the absence of WBFH results in an accuracy drop to 94.98%, highlighting its effectiveness in addressing class imbalance by focusing more on hard-to-classify samples. Thus, each component of the SLiTRANet model synergistically contributes to superior performance, demonstrating its technical superiority over traditional methodologies in complex classification tasks.

J. PERFORMANCE COMPARISON ANALYSIS OF THE PROPOSED MODEL WITH THE PRE-TRAINED CNN AND HYBRID-CNN ARCHITECTURES

A performance comparison analysis was done between the existing pre-trained CNN models and the proposed Transformer model. We have implemented prominent pre-trained CNN models such as VGG16/19, ResNet101/152, DenseNet121, etc. The findings demonstrated notable

TABLE 8. Ablation study of SLiTRANet: Performance impact of removing layers.

Model Configuration	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)
Without S-transform (preprocessed data input)	90.28	89.74	90.12	90.06
Without LGCN (Replaced with CNN)	84.72	83.48	84.38	84.57
Transformer (baseline model)	85.65	85.35	85.24	85.12
Without Transformer	89.38	87.23	89.14	89.02
Without Weighted Focal Binary Hinge Loss	94.98	95.20	94.90	94.30
SLiTRANet (Full Model)	**99.92**	**99.97**	**99.95**	**99.90**

TABLE 9. Performance comparison of transformer with existing CNN models for classification.

Model	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)
VGG16	90.80	88.87	91.10	90.42
VGG19	93.90	93.70	93.52	94.28
ResNet101	95.91	95.83	94.65	96.90
ResNet152	98.95	98.90	98.58	99.23
DenseNet121	99.04	98.80	98.24	99.60
CNN-LSTM	96.47	96.45	97.04	95.80
CNN-GRU	95.66	95.58	96.25	94.84
CNN-BiGRU	95.86	95.77	95.68	95.24
Proposed Model	**99.92**	**99.97**	**99.95**	**99.90**

improvements in precision and accuracy across a range of models. The VGG16 model obtained 90.80% accuracy and 88.87% precision, while the VGG19 model achieved 93.90% accuracy and 93.70% precision. ResNet101 achieved even more progress, attaining a precision of 95.83% and an accuracy of 95.91%. Surprisingly, ResNet152 revealed significant enhancement with 98.95% accuracy and 98.90% precision. DenseNet121 outperformed the competition, achieving 99.04% accuracy and 99.60% sensitivity. But the proposed Transformer model outscored them all, with a precision and accuracy combination of 99.97% and 99.92% respectively. Numerous variables contribute to the Transformer model's superiority over current CNN models in data classification applications. In processing data, where spatial interactions are intricate and non-local, transformers are indispensable for capturing long-range dependencies inside sequences. Transformers' second advantage is its self-attention mechanism, which facilitates thorough feature extraction and improves comprehension of the complex patterns. Transformers are also very parallelizable, which makes it possible to train them effectively on big datasets. This is useful for jobs like data classification that demand a lot of process power. Overall, the Transformer performs better in this area due to its capacity to process vast amounts of data, precisely model intricate spatial relationships, and extract significant characteristics. These results highlight the Transformer model's effectiveness and highlight its potential in dealing with challenging classification tasks. Table 9 summarizes the better classification results obtained by the proposed and existing CNN models.

Furthermore, we have also compared three hybrid CNN architectures, i.e., CNN-LSTM, CNN-GRU, and CNN-BiGRU, with our proposed Transformer architecture in Table 9. The Transformer model delivers remarkable results with an accuracy of 99.92%, precision of 99.97%, specificity of 99.95%, and sensitivity of 99.90%. On the other hand, CNN-LSTM, CNN-GRU, and CNN-BiGRU perform marginally worse, with respective accuracy rates of 96.47%, 95.66%, and 95.86%. From here, we can conclude that the proposed Transformer design architecture is quite appropriate for capturing depression-related minute EEG features, resulting in high MDD detection accuracy as demonstrated by the above analysis.

K. COMPARISON OF PROPOSED DEPRESSION DETECTION FRAMEWORK WITH THE EXISTING METHODS

In this section, we have presented a comparative performance analysis of the proposed approach with similar recent works reported in the literature. In order to have a fair comparison, we have chosen existing works evaluated on the same dataset. Table 10 depicts the performance comparison among different existing approaches for depression detection in terms of precision, accuracy, sensitivity, and specificity. With an accuracy of 99.92% and sensitivity, specificity, and precision of 99.97%, 99.95%, and 99.90%, respectively, our suggested model, which combines the S-Transform with the LGCN-Transformer, outperforms all other existing methods, demonstrating its better performance in depression identification. Mumtaz et al. [2] have employed power and symmetry features in conjunction with SVM to achieve perfect specificity and 98.4% accuracy. By combining alpha power, RWE, and MLPNN, Mahato and Paul [3] were able to achieve 93.33% accuracy with a good sensitivity but a lesser specificity. Non-linear characteristics were used with RBFSVM by Sharma et al. [4] to achieve good accuracy and balanced sensitivity and specificity. Similarly, Sharma et al. achieved near-perfect accuracy as well as good sensitivity and specificity by combining STFT with CNN-LSTM. Aydemir et al. [7] used SVM to apply melamine patterns, and while sensitivity was somewhat reduced, accuracy and specificity were quite high. Wavelet coherence was used by Khan et al. [8] to achieve reasonable accuracy with a balanced sensitivity and specificity. By using 2D-CNN on asymmetric pictures, Kang et al. [14] achieved excellent accuracy while maintaining a balanced sensitivity and specificity. A competitive performance was achieved by a CNN model proposed by Loh et al. [15], whose MDD performance is very close to our approach. The existing methods obtain satisfactory performance however, remain marginally lower

TABLE 10. Comparison table of our proposed architecture with the existing works.

Methods	Techniques Used	No. of Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
W. Mumtaz et al. [2]	Power and symmetry features + SVM	30 Normal, 34 Depressed	98.4	96.66	100	-
S. Mahato et al. [3]	Alpha power + RWE + MLPNN	30 Normal, 34 Depressed	93.33	94.44	87.78	-
G. Sharma et al. [4]	Non-linear features + RBFSVM	30 Normal, 34 Depressed	98.9	99.2	99.7	-
G. Sharma et al. [5]	STFT + CNN-LSTM	30 Normal, 34 Depressed	99.9	100	99.81	99.47
E. Aydemir [7]	Melamine pattern + SVM	30 Normal, 34 Depressed	99.1	98.4	99.8	99.8
D.M.Khan et al. [8]	Wavelet coherence	30 Normal, 34 Depressed	98.1	98.0	99.82	-
M. Kang et al. [14]	Asymmetry image + 2D-CNN	30 Normal, 34 Depressed	98.8	99.1	98.5	-
H.W.Loh et al. [15]	CNN	30 Normal, 34 Depressed	99.58	99.70	99.48	99.40
Proposed Model (HUSM dataset)	S Transform + LGCN-Transformer	30 Normal, 34 Depressed	**99.92**	**99.97**	**99.95**	**99.90**
Proposed Model (MODMA dataset)	S Transform + LGCN-Transformer	29 Normal, 24 Depressed	99.68	99.65	99.70	99.75

than the depression detection performance obtained by the proposed method.

V. CONCLUSION

This work proposed a smart IoMT-based end-to-end framework for automated real-time MDD diagnosis and remote monitoring. This research initiative seeks to advance the field of IoMT-enabled smart healthcare systems for remote monitoring of the mental state of a person through wearable EEG headsets and cutting-edge deep learning approaches. This work has proposed a method for depression identification by integrating the S-transform with the modified Linear Graph Convolution Network (LGCN) and Transformer-based models. This method achieved significantly high depression detection performance with 99.92% accuracy, 99.90% sensitivity, 99.95% specificity, and 99.97% precision, highlighting the effectiveness of the proposed framework. Compared to other existing methods, such as LGCN combined with traditional CNNs, the Stockwell Transform-modified LGCN+Transformer architecture demonstrates its superiority in capturing complex temporal and spatial correlations within depression-related EEG data.

Future directions in depression detection include a wide range of approaches targeted at improving the precision, scalability, and usefulness of current techniques. Integrating multimodal data sources, such as physiological signs, social media activity, and EHRs, are possible future avenues. Researchers can learn more about the intricate interactions between biological, psychological, and environmental components that contribute to depression by combining these various data sources. The creation of interpretable deep-learning models that can forecast depression and shed light on the underlying processes causing depressive symptoms is another area of future research. Methods like explainable AI have the potential to clarify the key elements that lead to the development and progression of depression. Furthermore, developments in wearable sensor technology have the potential for ongoing observation of behavioral and physiological indicators linked to depression, facilitating early identification and customized therapies. Lastly, developing predictive tactics and tailored interventions can be aided by using longitudinal data and deep learning approaches to find trends in the trajectory of depression symptoms over time.

ACKNOWLEDGMENT

The authors express their sincere gratitude to Vellore Institute of Technology for providing the necessary resources and facilities to carry out this study.

REFERENCES

[1] Inst. Health Metrics Eval. *Global Health Data Exchange (GHDx)*. Accessed: Mar. 24, 2024. [Online]. Available: https://vizhub.healthdata.org/gbd-results/

[2] W. Mumtaz, L. Xia, S. S. A. Ali, M. A. M. Yasin, M. Hussain, and A. S. Malik, "Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD)," *Biomed. Signal Process. Control*, vol. 31, pp. 108–115, Jan. 2017.

[3] S. Mahato and S. Paul, "Detection of major depressive disorder using linear and non-linear features from EEG signals," *Microsyst. Technol.*, vol. 25, no. 3, pp. 1065–1076, Mar. 2019.

[4] G. Sharma, A. M. Joshi, and E. S. Pilli, "DepML: An efficient machine learning-based MDD detection system in IoMT framework," *Social Netw. Comput. Sci.*, vol. 3, no. 5, p. 394, Jul. 2022.

[5] G. Sharma, A. M. Joshi, R. Gupta, and L. R. Cenkeramaddi, "DepCap: A smart healthcare framework for EEG based depression detection using time-frequency response and deep neural network," *IEEE Access*, vol. 11, pp. 52327–52338, 2023.

[6] M. Kang, H. Kwon, J.-H. Park, S. Kang, and Y. Lee, "Deep-asymmetry: Asymmetry matrix image for deep learning method in pre-screening depression," *Sensors*, vol. 20, no. 22, p. 6526, Nov. 2020.

[7] E. Aydemir, T. Tuncer, S. Dogan, R. Gururajan, and U. R. Acharya, "Automated major depressive disorder detection using melamine pattern with EEG signals," *Int. J. Speech Technol.*, vol. 51, no. 9, pp. 6449–6466, Sep. 2021.

[8] D. M. Khan, K. Masroor, M. F. M. Jailani, N. Yahya, M. Z. Yusoff, and S. M. Khan, "Development of wavelet coherence EEG as a biomarker for diagnosis of major depressive disorder," *IEEE Sensors J.*, vol. 22, no. 5, pp. 4315–4325, Mar. 2022.

[9] S. De, P. Mukherjee, D. Biswas, and A. Halder Roy, "A fusion of CNN and grey wolf optimization-enhanced BiGRU for epileptic seizure recognition using EEG signals," in *Proc. 3rd Int. Conf. Mobile Netw. Wireless Commun. (ICMNWC)*, Dec. 2023, pp. 1–6.

[10] G. Sharma, A. Parashar, and A. M. Joshi, "DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102393.

[11] S. De, P. Mukherjee, D. Konar, and A. H. Roy, "AromaNet: Integrating attention mechanism with convolutional neural network for olfactory perception classification using EEG signals," in *Proc. 4th Int. Conf. Commun., Comput. Ind. 6.0*, Dec. 2023, pp. 1–6.

[12] S. De, P. Mukherjee, and A. H. Roy, "EEG-based intelligence quotient assessment using 1D convolutional neural network," in *Proc. 8th Int. Conf. Comput. Devices Commun. (CODEC)*, Dec. 2023, pp. 1–2.

[13] Z. Wan, J. Huang, H. Zhang, H. Zhou, J. Yang, and N. Zhong, "HybridEEGNet: A convolutional neural network for EEG feature learning and depression discrimination," *IEEE Access*, vol. 8, pp. 30332–30342, 2020.

[14] M. Kang, J. Park, S. Kang, and Y. Lee, "Low channel electroencephalogram based deep learning method to pre-screening depression," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 449–451.

[15] H. W. Loh, C. P. Ooi, E. Aydemir, T. Tuncer, S. Dogan, and U. R. Acharya, "Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals," *Expert Syst.*, vol. 39, no. 3, Mar. 2022, Art. no. e12773.

[16] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. Herrera-Viedma, and O. Krejcar, "DeprNet: A deep convolution neural network framework for detecting depression using EEG," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.

[17] S. De, P. Mukherjee, and A. H. Roy, "A novel deep learning-based approach for hypertension level detection using PPG," in *Proc. IEEE Silchar Subsection Conf. (SILCON)*, Silchar, India, Nov. 2023, pp. 1–6.

[18] S. De, P. Mukherjee, D. Konar, and A. H. Roy, "EEG-based taste perception classification using PCA enhanced attention-TLSTM neural network," in *Proc. 4th Int. Conf. Commun., Comput. Ind. 6.0*, Bangalore, India, Dec. 2023, pp. 1–6.

[19] A. Saeedi, M. Saeedi, A. Maghsoudi, and A. Shalbaf, "Major depressive disorder diagnosis based on effective connectivity in EEG signals: A convolutional neural network and long short-term memory approach," *Cognit. Neurodyn.*, vol. 15, no. 2, pp. 239–252, Apr. 2021.

[20] S. De, P. Mukherjee, and A. H. Roy, "A hybrid pain assessment approach with stacked autoencoders and attention-based CP-LSTM," in *Proc. Int. Conf. Ambient Intell., Knowl. Informat. Ind. Electron. (AIKIIE)*, Ballari, India, Nov. 2023, pp. 1–6.

[21] D. M. Khan, N. Yahya, N. Kamel, and I. Faye, "Automated diagnosis of major depressive disorder using brain effective connectivity and 3D convolutional neural network," *IEEE Access*, vol. 9, pp. 8835–8846, 2021.

[22] D. Konar, S. De, P. Mukherjee, and A. H. Roy, "A novel human stress level detection technique using EEG," in *Proc. Int. Conf. Netw., Multimedia Inf. Technol. (NMITCON)*, Bengaluru, India, Sep. 2023, pp. 1–6.

[23] Y. Zhao, C. Dong, G. Zhang, Y. Wang, X. Chen, W. Jia, Q. Yuan, F. Xu, and Y. Zheng, "EEG-based seizure detection using linear graph convolution network with focal loss," *Comput. Methods Programs Biomed.*, vol. 208, Sep. 2021, Art. no. 106277.

[24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.* Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 6000–6010.

[25] U. P. Shukla and S. Garg, "LightFFNet: MDD prediction on EEG quantitative biomarkers," in *Proc. Int. Conf. Eng. Emerg. Technol. (ICEET)*, Kuala Lumpur, Malaysia, Oct. 2022, pp. 1–6.

[26] H. Cai et al., "A multi-modal open dataset for mental-disorder analysis," *Sci. Data*, vol. 9, no. 1, p. 178, Apr. 2022.

[27] S. De and A. K. Gupta, "A quantum machine learning framework for driver drowsiness detection using biopotential signals and head movement analysis," in *Proc. IEEE Int. Conf. Women Innov., Technol. Entrepreneurship (ICWITE)*, Bangalore, India, Feb. 2024, pp. 461–466.

[28] H. Cheng, Z. Wang, W. Liu, and C. Hu, "Depression classification using log-Mel energy based on EEG," in *Proc. 3rd Int. Conf. Frontiers Electron., Inf. Comput. Technol. (ICFEICT)*, Yangzhou, China, May 2023, pp. 224–229.

[29] P. Jain, A. M. Joshi, and S. P. Mohanty, "iGLU 1.1: Towards a glucose-insulin model based closed loop IoMT framework for automatic insulin control of diabetic patients," in *Proc. IEEE 6th World Forum Internet Things (WF-IoT)*, Jun. 2020, pp. 1–6.

[30] S. Lee, P. Huang, M. Liang, J. Hong, and J. Chen, "Development of anarrhythmia monitoring system and human study," *IEEE Trans. Consum. Electron.*, vol. 64, no. 4, pp. 442–451, Nov. 2018.

[31] N. Dey, A. S. Ashour, F. Shi, S. J. Fong, and R. S. Sherratt, "Developing residential wireless sensor networks for ECG healthcare monitoring," *IEEE Trans. Consum. Electron.*, vol. 63, no. 4, pp. 442–449, Nov. 2017.

[32] S. De, A. Sayyad, H. Kotian, and A. K. Gupta, "ParViT: A modified vision transformer architecture for Parkinson's disease identification using EEG signals," in *Proc. Int. Conf. Smart Syst. Appl. Electr. Sci. (ICSSES)*, Tumakuru, India, May 2024, pp. 1–6.

[33] S. De, P. Mukherjee, and A. H. Roy, "BPWave: An advanced deep learning framework for continuous non-invasive cuffless blood pressure estimation and hypertension detection," in *Proc. Int. Conf. Smart Syst. Appl. Electr. Sci. (ICSSES)*, Tumakuru, India, May 2024, pp. 1–6.

[34] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "Neuro-detect: A machine learning-based fast and accurate seizure detection system in the IoMT," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 359–368, Aug. 2019.

[35] L. Lei, Z. Liu, Y. Zhang, M. Guo, P. Liu, X. Hu, C. Yang, A. Zhang, N. Sun, Y. Wang, and K. Zhang, "EEG microstates as markers of major depressive disorder and predictors of response to SSRIs therapy," *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, vol. 116, Jun. 2022, Art. no. 110514.

[36] B. C. Kavanaugh, A. M. Fukuda, Z. T. Gemelli, R. Thorpe, E. Tirrell, M. Vigne, S. R. Jones, and L. L. Carpenter, "Pre-treatment frontal beta events are associated with executive dysfunction improvement after repetitive transcranial magnetic stimulation for depression: A preliminary report," *J. Psychiatric Res.*, vol. 168, pp. 71–81, Dec. 2023.

[37] S. Yasin, S. A. Hussain, S. Aslan, I. Raza, M. Muzammel, and A. Othmani, "EEG based major depressive disorder and bipolar disorder detection using neural networks: A review," *Comput. Methods Programs Biomed.*, vol. 202, Apr. 2021, Art. no. 106007.

[38] D. Watts, R. F. Pulice, J. Reilly, A. R. Brunoni, F. Kapczinski, and I. C. Passos, "Predicting treatment response using EEG in major depressive disorder: A machine-learning meta-analysis," *Transl. Psychiatry*, vol. 12, no. 1, p. 332, Aug. 2022.

[39] M. M. Eid, W. Yundong, G. B. Mensah, and P. Pudasaini, "Treating psychological depression utilising artificial intelligence: AI for precision medicine- focus on procedures," *Mesopotamian J. Artif. Intell. Healthcare*, vol. 2023, pp. 76–81, Dec. 2023.

[40] S. Assous and B. Boashash, "Evaluation of the modified S-transform for time-frequency synchrony analysis and source localisation," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 49, Feb. 2012.

[41] S. A. Akar, S. Kara, S. Agambayev, and V. Bilgiç, "Nonlinear analysis of EEGs of patients with major depression during different emotional states," *Comput. Biol. Med.*, vol. 67, pp. 49–60, Dec. 2015.

[42] X. Li, B. Hu, S. Sun, and H. Cai, "EEG-based mild depressive detection using feature selection methods and classifiers," *Comput. Methods Programs Biomed.*, vol. 136, pp. 151–161, Nov. 2016.

[43] L. Simmatis, E. E. Russo, J. Geraci, I. E. Harmsen, and N. Samuel, "Technical and clinical considerations for electroencephalography-based biomarkers for major depressive disorder," *NPJ Mental Health Res.*, vol. 2, no. 1, p. 18, Oct. 2023.

[44] C.-T. Wu, H.-C. Huang, S. Huang, I.-M. Chen, S.-C. Liao, C.-K. Chen, C. Lin, S.-H. Lee, M.-H. Chen, C.-F. Tsai, C.-H. Weng, L.-W. Ko, T.-P. Jung, and Y.-H. Liu, "Resting-state EEG signal for major depressive disorder detection: A systematic validation on a large and diverse dataset," *Biosensors*, vol. 11, no. 12, p. 499, Dec. 2021.

[45] N. Pusarla, A. Singh, and S. Tripathi, "Normal inverse Gaussian features for EEG-based automatic emotion recognition," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 6503111.

[46] N. Pusarla, A. Singh, and S. Tripathi, "Learning DenseNet features from EEG based spectrograms for subject independent emotion recognition," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103485.

SAGNIK DE (Student Member, IEEE) is currently pursuing the Bachelor of Technology degree in electronics and communication engineering with the Institute of Radio Physics and Electronics, University of Calcutta, Kolkata. As a dedicated and ambitious undergraduate student, he is deeply engaged in academic research, with his work being presented at IEEE conferences and submitted to prestigious journals. His research interests include artificial intelligence and the IoT in healthcare, human–computer interaction, computational neuroscience, biomedical signal processing, and medical image analysis.

ANURAG SINGH (Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Mahatma Jyotiba Phule Rohilkhand University, Bareilly, India, in 2009, the M.Tech. degree in electronics and communication engineering from the Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India, in 2012, and the Ph.D. degree in biomedical signal processing from the Indian Institute of Technology Guwahati, Guwahati, India, in 2017. He is currently an Associate Professor with the Electronics and Communication Engineering Department, International Institute of Information Technology Naya Raipur (IIIT-NR), Naya Raipur, India. Prior to joining IIIT-NR, he was associated as a Faculty Member with MNIT Jaipur and Thapar University, Patiala, Punjab. He was a part of the Delegation for the Indo-Japan Student Exchange Program JENESYS-2011. He has published several articles in journals and conference proceedings of international repute. He is serving as a peer reviewer for many reputed international journals and conferences, including IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, IEEE ACCESS, IEEE SENSORS JOURNAL, *Biomedical Signal Processing and Control* (Elsevier), *Computers and Electrical Engineering* (Elsevier), *International Journal of Machine Learning and Cybernetics* (Springer), *International Journal of Electronics Letters* (Taylor & Francis), and *Frontiers in Physiology*. His research interests include machine learning, AI and the IoT in healthcare, biomedical signal and image processing, and brain–computer interfacing (BCI). He has been an active member of the Signal Processing Society and the Engineering in Medicine and Biology Society. He was a recipient of the International Travel Grant (Young Scientist) by the Department of Science and Technology, Government of India.

VIVEK TIWARI (Senior Member, IEEE) was associated as a Professor-In-Charge with the Department of Data Science and AI, IIIT Naya Raipur, Chhattisgarh, India. He was a Faculty Member with the Department of Computer Science and Engineering, IIIT-Naya Raipur, from June 2016 to June 2023. Since June 2023, he has been an Associate Professor with the Computer Science and Engineering Department, ABV-Indian Institute of Information Technology and Management (ABV-IIITM), Gwalior, where he is currently an Associate Dean (Academics). His current research interests include machine/deep learning, data mining, pattern recognition, business analytics, data warehousing, activities recognition, surveillance analytics, human–computer interaction, and computer vision (application).

HARSHITA PATEL received the Ph.D. degree from the Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India, in 2017. She is currently an Associate Professor with the School of Computer Science Engineering and Information Systems (SCORE), Vellore Institute of Technology, Vellore, India. Additionally, she has qualified for national-level competitive exams, such as GATE and UGC-NET. She has more than 14 years of teaching and research experience. She has published a good number of research papers in various conferences and journals of international repute which are indexed in SCOPUS and SCI. She has also delivered expert lectures in reputed workshops. Her research interests include data mining, machine learning, and artificial intelligence.

G. N. VIVEKANANDA (Senior Member, IEEE) is currently an Associate Professor with the School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore. He acted as a Resource Person for more than 85 programs. He is a Certified Ethical Hacker, Microsoft Certified Educator, Certified Professional, and IBM Certified Academic Associate. He published many articles in peer-reviewed journals. His research interests include AI and ML, networks, security, the IoT, and blockchain technologies. He is a Life Member of various prestigious organizations, such as IEI, CSI, and ISTE. He was a recipient of the IEI Young Engineers Award and the CSI Young Researcher Award. He chaired sessions and was guest of honor at the conferences.

DHARMENDRA SINGH RAJPUT received the Ph.D. degree from NIT, Bhopal, India, in 2014. Since June 2014, he has been a Professor with the Department of Software and Systems Engineering, SCORE, Vellore Institute of Technology, Vellore. He is an ICT Researcher. He is involved in the Erasmus + Program of the European Union-funded project of nine crores which is received from the University of Nottingham, U.K. He has visited eight countries namely U.K., France, and Singapore, for academic purposes. He has published more than 40 reputed journal articles, five edited books published under reputed publishers, and 20 papers presented at a reputed international conference. His research expertise is machine learning and sentiment analysis. His research interests include machine learning, soft computing, and big data analytics. He is also a guest and academic editor of various reputed journals.