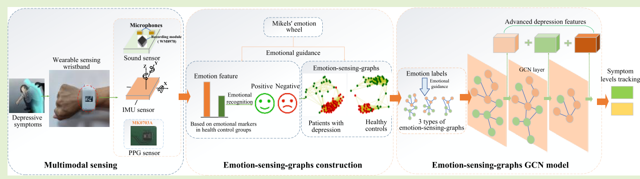
IEEE Sensors Council IEEE SENSORS JOURNAL, VOL. 24, NO. 3, 1 FEBRUARY 2024

Using Wearable and Structured Emotion-Sensing-Graphs for Assessment of Depressive Symptoms in Patients Undergoing Treatment

Daili Yang, Yunge Li, Bin Gao, *Senior Member, IEEE*, Wai Lok Woo, *Senior Member, IEEE*, Yiwei Zhang, Keith M. Kendrick, and Lizhu Luo

***Abstract*—Depression, as a common mental illness, has become a significant public health issue, and the recurrence rate for patients with depression who have been treated is relatively high. In this study, a mental health monitoring system based on wearable sensing wristbands with sensors for voice, activity, and heart rate has been developed. Using this system, we perform a therapeutic monitoring study for hospitalized patients with depression and healthy controls to investigate multimodal changes before, during, and after a course of treatment. The obtained results demonstrate that there are significant changes in multimodal features such as audio short-time energy and angular velocity shape skewness with the remission of depressive symptoms. According to Mikels' emotion wheel, a day's data for subjects is defined as three types of emotional units and the emotional state of each emotional unit is recognized as positive or negative emotions. With this, emotion-sensing-graphs guided by Mikels' emotion wheel theory are constructed. The analysis of emotion-sensing-graphs reveals that the same emotions are more closely linked to each other and the average degree and proportion of positive emotion nodes after a course of treatment have increased significantly. Finally, an emotion-sensing-graph graph convolutional network (ESG-GCN) model fused three types of emotion-sensing-graphs with emotion labels has been developed to assess the levels of depression, thereby monitoring the changes in depressive symptoms. Compared with classical machine learning models, the accuracy, F1 score, and recall rate of the model perform best and the model achieves a verification accuracy of 0.83.**



***Index Terms*—Depression, emotion-sensing-graph graph convolutional network (ESG-GCN) model, emotion-sensing-graphs, Mikels' emotion wheel, wearable sensing wristbands.**

Manuscript received 28 September 2023; accepted 28 November 2023. Date of publication 12 December 2023; date of current version 31 January 2024. The work of Lizhu Luo was supported in part by the China Postdoctoral Science Foundation under Grant 2021T140094 and in part by the National Natural Science Foundation of China under Grant 31800961. The associate editor coordinating the review of this article and approving it for publication was Dr. Edward Sazonov. *(Corresponding authors: Bin Gao; Lizhu Luo.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Committee of the University of Electronic Science and Technology of China.

Daili Yang is with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China, and also with the School of Intelligent Manufacturing, Sichuan University of Arts and Science, Dazhou 635000, China.

Yunge Li, Keith M. Kendrick, and Lizhu Luo are with the Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: luolizhu@uestc.edu.cn).

Bin Gao and Yiwei Zhang are with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: bin_gao@uestc.edu.cn).

Wai Lok Woo is with the Department of Computer and Information Sciences, Northumbria University, NE1 8ST Newcastle upon Tyne, U.K.

Digital Object Identifier 10.1109/JSEN.2023.3339498

I. Introduction

DEPRESSION, as a common mental illness, has become a global public health problem. According to the World Health Organization (WHO), more than 300 million people currently suffer from depression [1]. Long-lasting depression can cause a person to fall into a long-term negative mood, which can result in a lack of self-confidence, feelings of guilt, or even a loss of interest in life [2]. Depression can be alleviated or even cured by psychological therapies, antidepressant medications, or a combination of these approaches [3]. Moreover, as people become more aware of and attach importance to depression, many patients with depression can also seek help from professional psychiatrists in a timely manner. However, for patients who have achieved remission after a course of treatment, depression can easily recur. Research has shown that the recurrence rate of depression within three years is as high as 70%–80% [4], [5]. In addition, the diagnosis of depression is mainly determined by professional psychiatrists inquiring about the patient's condition and using criteria defined in clinical assessments (such as DSM-IV or DSM-V) and scales (such as Hamilton depression (HAMD) scale) so that patients and their families have limited insight into changes in depression symptoms. Clinicians are also unlikely to assess changes in a patient's condition unless they report regularly for consultations, making it difficult for patients to receive timely medical assistance as soon as they experience a relapse. Therefore, the establishment of long-term dynamic monitoring and evaluation platform for depression based on wearable sensing technology may be an effective way to detect changes in depressive symptoms, thereby providing early warning and preventing further deterioration.

Generally, patients with depression exhibit mental retardation, decreased volitional activity, as well as low mood and emotional abnormalities. In depressed patients, these characteristics can be sensed through voice [6], [7], activity [8], [9], and physiological [10] signals. Wang et al. [11] extracted three types of features of Mel-scale frequency cepstral coefficients (MFCCs), short-term energy, and spectral entropy of speech signals reflecting depression information based on audio difference normalization algorithm. Then, a depression recognition model based on convolution neural network (CNN) and generative antagonism network (GAN) was constructed. The results showed that the depression recognition error was reduced on the AViD-Corpus and DAIC-WOZ datasets, and the RMSE and MAE values were increased by more than 5% compared to existing methods. Seal et al. [12] proposed a deep learning (DL)-based convolutional neural network (CNN) model, called DeprNet, for identifying depressed and healthy subjects based on the EEG data. The model achieved accuracies of 0.9937 and 0.914 in two experiments, namely, the recordwise split and the subjectwise split. The results showed that the performance of DeprNet was superior to the other eight baseline models. Therefore, mental health monitoring systems based on wearable sensing devices can objectively reflect changes in symptoms of patients with depression by collecting information from sensors such as voice, activity, and heart rate [13], [14], [15]. Many previous research efforts have shown that this is a noninvasive, objective, and effective way to track and identify depression symptoms. A wearable device was utilized to collect multimodal information such as body movement, steps, skin temperature, heart rate, and sleep time from patients with depression and healthy controls, and an evaluation model for depression symptom severity was established using machine learning [16]. Pedrelli et al. [17] have validated that it was effective and feasible to assess depressive symptom levels through behavioral and physiological features, which were collected from wearable wristbands and smartphones. Rui et al. [18] have built a predictive model based on symptom features that were exploited from multimodal data collected by wearable sensors and mobile phones to continuously assess whether or not students were depressed. Thus, through long-term monitoring and analysis of multimodal features from wearable sensing devices, it is possible for patients and doctors to observe depressive symptoms more reliably and for doctors to make unscheduled evaluations or change the course of treatment in a more personalized manner.

Graph networks not only contain the properties of things but also describe the intrinsic connections among things. Currently, a number of studies have focused on digging deeper into the information related to depression to construct graph networks and then using graph neural networks (GNNs) to evaluate and identify depression levels. Bidja [19] has found node features and edge metrics highly correlated with depression and built graph networks that accurately represented the depression dataset, achieving over 80% accuracy in predicting depression using GNN on smartphones and wearable sensors. Chen et al. [20] formulated the subject's EEG signal as a graph structure in which the edges were constructed by a combination of local and global connections. Based on this, a self-focused graph pooling module was introduced to build a self-attention graph pooling with soft label (SGP-SL) model for the detection of major depressive disorder (MDD). Therefore, it is effective and feasible to fully exploit intrinsic information in depression to construct a graph structure for the prediction and assessment of depression. However, depression is a mood disorder, and most of studies have endeavored to find the relationship in the data itself, while few studies integrated psychological prior knowledge such as emotion theory into graph network construction based on wearable sensing devices data.

Majority of prior studies have focused on the identification, assessment, and prediction of depression based on multimodal sensors, while dynamic monitoring of depressive symptoms in patients with depression during treatment has rarely been investigated. In addition, human emotions are crucial for identifying a person's behavior and mental state [21]. For patients with depression, the symptoms are prolonged with persistent low mood and anhedonia, and some patients with depression are also irritable and emotionally abnormal. These characteristics reflect a direct relationship between depression and emotion. In this study, we track the treatment process of hospitalized patients with depression using wearable sensing wristbands integrating voice, activity, and heart rate sensors. The multimodal signals of hospitalized patients with depression before and after a course of treatment are analyzed. A day's data of patients with depression and healthy controls are defined using different emotional units. According to the psychological prior knowledge of Mikels' emotional wheel [22], the emotion of each emotional unit is recognized as positive or negative emotion. Then, emotion-sensing-graphs guided by Mikels' emotion wheel theory are constructed. Meanwhile, the emotion-sensing-graphs between patients with depression before and after a course of treatment and healthy controls are compared. Finally, an emotion-sensing-graph graph convolutional network (ESG-GCN) model fusing three types of emotion-sensing-graphs with emotion labels is built to monitor the changes of depressive symptoms in patients.

II. METHODOLOGY

Fig. 1 shows the overall architecture of the designed platform. In stage 1, a multisensor wearable wristband has been developed. Fifty-seven patients with depression and 21 healthy controls who are evaluated by psychiatrists are selected and their basic information has been recorded. Subsequently, the baseline data of audio information for subjects are collected.

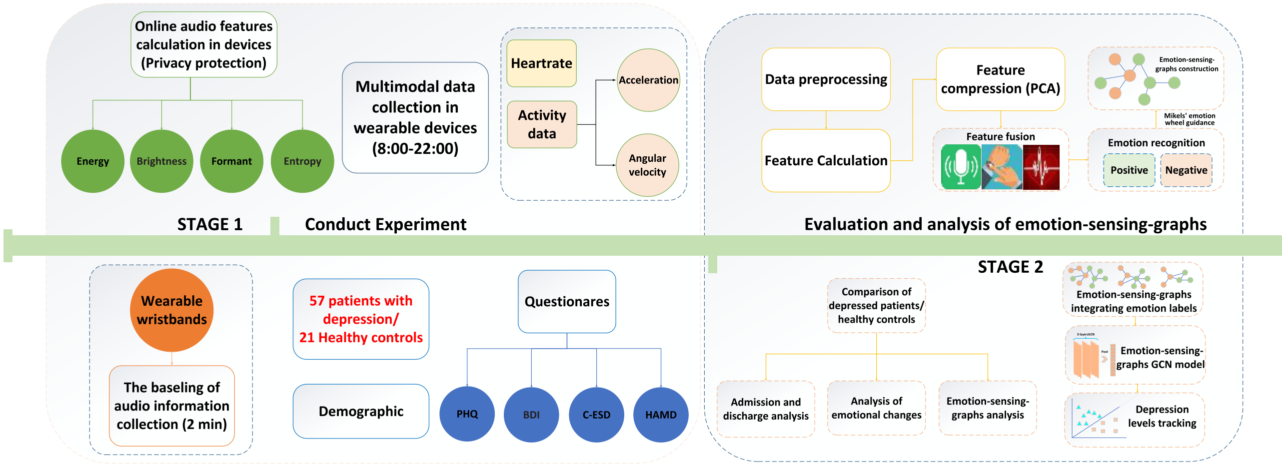


Fig. 1. Overall proposed architecture consisting of two stages.

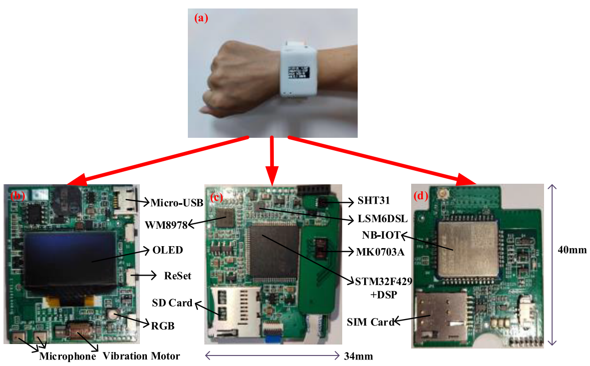


Fig. 2. Developed wearable sensing device. (a) Wearable sensing wristband. (b) Front view of a core board. (c) Rear view of a core board. (d) NB-IoT board.

During the experiment, the subjects wear a wearable sensing wristband for one day at a time (8:00 A.M.–22:00 P.M.) to collect the data of audio, activity, and heartrate, and they are asked to fill out questionnaires to assess their psychological states. In stage 2, activity and heartrate features are calculated after the multimodal data is preprocessed. We then use a principal component analysis (PCA) to compress and fuse multimodal features. The state of patients with depression is often directly related to their moods. Therefore, considering emotional changes and the relationship among emotions, we have separately defined 10, 20, and 30 min as emotional units to construct emotion-sensing-graphs guiding by Mikels' emotion wheel. Thus, we have performed admission and discharge analysis, analysis of emotional changes, and analysis of emotion-sensing-graphs to observe changes for patients with depression after a course of treatment. Finally, an ESG-GCN model has been developed to monitor and track changes in depressive symptom levels for patients with depression and three types of emotion-sensing-graphs with emotion labels are fused to train and improve the performances of the model.

A. Wearable Sensing Wristbands

We have developed a wearable sensing wristband that incorporates a diverse range of sensors, enabling noninvasive and objective collection of voice, activity, and physiological information over a long period of time. The diagrammatic representation of a wearable sensing wristband can be observed in Fig. 2. In order to optimize the size of the wearable sensing wristbands, reduce power consumption, and clarify functional zoning, the wearable sensing wristbands are designed as a core board, a narrowband Internet of Things (NB-IoT) board, and a heart rate board. The board size is 34 $\times$ 40 mm, and the three boards adopt a multilayer overlapping structure inside the wearable sensing wristbands. Most sensors and devices are integrated on the core board, mainly including recording module WM8978, acceleration, gyroscope sensor LSM6DSL, and an OLED where the operating status and real-time sensor data are displayed. In order to accomplish real-time transmission of data, the NB-IoT board encompasses an NB-IoT module. The wearer's heart rate and body temperature are collected by the MK0703A and SHT31 sensors on the heart rate board, respectively. The entire system runs in an orderly manner under the control of the main controller STM32F405G, which is based on the arm-cortex4 kernel and integrates a DSP module. In addition, a Micro-USB interface and SD card module are utilized for program debugging and data storage, respectively. The whole system is powered by a lithium battery (3.7 V, 1800 mAh) and supplies power to each module through a voltage conditioning circuit.

B. Feature Extraction of Wearable Sensing Wristband Data

Audio, activity, and heart rate data have been gathered from the newly designed wearable sensing wristbands. However, due to the presence of noise and external environmental interference in the original signals, direct utilization of the signals for analysis and processing is challenging. Therefore, feature extraction is carried out using the collected data from the wearable sensing wristbands for subsequent analysis and building models.

1) Audio Feature Extraction: The audio processing flow is shown in Fig. 3. In the field of speech signal processing, the preprocessing techniques for speech signals collected in a natural environment mainly include digital discretization, framing, preemphasis, and windowing. The WM8978 encoding and decoding chip for wearable sensing wristbands has completed the digitization of the audio signal using an internal ADC converter, and its sampling frequency is set

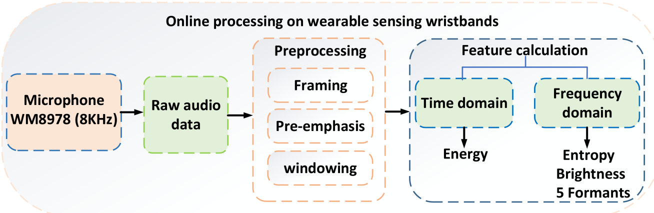


Fig. 3. Audio data processing flow on wearable sensing wristbands.

to 8 KHz to ensure that the audio signal is not distorted. In addition, preprocessing methods, such as framing, preemphasis, and windowing, are embedded into wearable sensing wristbands. Considering the limited computing power of the devices and privacy protection concern, only eight audio features of short-term energy, entropy, formants (including five values), and brightness [23], [24] are extracted online and real time on wearable sensing wristbands.

2) Activity Features Extraction: We have collected activity data from wearable sensing wristbands, with a sampling frequency of 85 Hz. The activity data are mainly generated through wrist motion, which includes the 3-D spatial data of the three-axis gyroscope and the three-axis accelerometer. To facilitate feature extraction, we use the synthetic acceleration and synthetic angular velocity scheme

$$a_i = \sqrt{\left(a_i^x\right)^2 + \left(a_i^y\right)^2 + \left(a_i^z\right)^2} \tag{1}$$

$$q_i = \sqrt{\left(q_i^x\right)^2 + \left(q_i^y\right)^2 + \left(q_i^z\right)^2} \tag{2}$$

where a_i and q_i are the synthetic acceleration and synthetic angular velocity at the ith moment, respectively, and a_i^x, a_i^y, and a_i^z represent triaxial (x-, y-, and z-axes) accelerometer readings. q_i^x, q_i^y, and q_i^z represent triaxial (x-, y-, and z-axes) angular velocity readings. The original data of synthetic acceleration and synthetic angular velocity are processed by sliding windows, and each window contains 256 sampling points. To better capture the activity state and remove abnormal data, adjacent windows overlap by 50%. From the data within each window, time-domain features are computed and frequency-domain features are simultaneously extracted from the transformed frequency-domain sequences using fast Fourier transform (FFT). We have also calculated the signal amplitude area (SMA), which is the sum of the areas enclosed by the acceleration values of the three axes (x, y, z). Details of the features in the time and frequency domains are shown in Table I. The shape feature value is the 2-D area formed by FFT results. The shape statistics are defined as follows:

$$u_{\text{shape}} = \frac{1}{S}\sum_{i=1}^{N} iC(i) \tag{3}$$

$$\sigma_{\text{shape}} = \sqrt{\frac{1}{S}\sum_{i=1}^{N}\left(i - u_{\text{shape}}\right)^2 C(i)} \tag{4}$$

$$\gamma_{\text{shape}} = \frac{1}{S}\sum_{i=1}^{N}\left(\frac{i - u_{\text{shape}}}{\sigma_{\text{shpae}}}\right)^3 C(i) \tag{5}$$

where $S = \sum_{i=1}^{N} C(i)$, $C(i)$ is the power spectral density (PSD) magnitude for the ith frequency bin. In this article,

TABLE I
43 TIME- AND FREQUENCY-DOMAIN FEATURES OF ACTIVITY DATA

ID	Feature	Description
1	acce_time_mean, ang_time_mean	Average value of samples in a window
2	acce_time_var, ang_time_var	Standard deviation of samples
3	acce_time_std, ang_time_std	Minimum of samples in a window
4	acce_time_mode, ang_time_mode	Maximum of samples in a window
5	acce_time_max, ang_time_max	The value with the largest frequency
6	acce_time_min, ang_time_min	Variance of samples in a window
7	acce_time_range, ang_time_range	Maximum minus minimum
8	acce_time_over_zero, ang_time_over_zero	The number of more than the average
9	acce_time_sqrt, ang_time_sqrt	Root mean square
10	acce_fft_dc, ang_fft_dc	Average value of samples in a window
11	acce_shape_mean, ang_shape_mean	
12	acce_shape_var, ang_shape_var	Mean, standard deviation, skewness, variance of shape about an FFT window
13	acce_shape_std, ang_shape_std	
14	acce_shape_skew, ang_shape_skew	
15	acce_fft_mean, ang_fft_mean	
16	acce_fft_var, ang_fft_var	
17	acce_fft_std, ang_fft_std	
18	acce_fft_skew, ang_fft_skew	Mean, variance, standard deviation, kurtosis, skewness, energy and entropy of an FFT window
19	acce_fft_kurt, ang_fft_kurt	
20	acce_fft_energy, ang_fft_energy	
21	acce_fft_entropy, ang_fft_entropy	
22	acce_sma	Signal Magnitude Area

Note: acce_time and acce_fft represent the prefix of the time and frequency domain features of acceleration. ang_time and ang_fft represent the prefix of the time and frequency domain features of angular velocity. acce_shape and ang_shape represents the prefix of shape statistical features of the power spectral density (PSD).

"ang_shape_skew" represents the shape skewness feature of the angular velocity.

3) Heartrate Feature Extraction: Changes of heart rate will reflect both altered physical and psychological states. Therefore, the calculation of heart rate is particularly critical. The MK0703A module for wearable sensing wristbands uses photoplethysmography (PPG) signals to calculate the heart rate value in real time. In this study, eight features are extracted from heart rate data, including mean, variance, standard deviation, mode, maximum, minimum, range, and first-order difference. The calculation of various features is illustrated as follows:

$$\text{mean} = \frac{1}{N}\sum_{n=1}^{N} S_n \tag{6}$$

$$\text{var} = \frac{1}{N-1}\sum_{n=1}^{N}\left(S_n - \text{mean}\right)^2 \tag{7}$$

$$\text{diff} = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|S_{n+1} - S_n\right| \tag{8}$$

$$\text{range} = \max - \min \tag{9}$$

where S_n represents the heart rate at nth time and N represents the length of a segment of data.

C. Emotion-Sensing-Graphs Model

The categories of emotional experience for major depressive patients include depressed mood, feeling sad, loss of pleasure, feeling empty, irritable mood, inappropriate guilt, and feelings of worthlessness in the DSM-IV [25]. According to Mikels' emotion wheel, there are eight basic emotions. Excitement, awe, contentment, and amusement belonged to the positive emotions are distributed in the left semicircle,

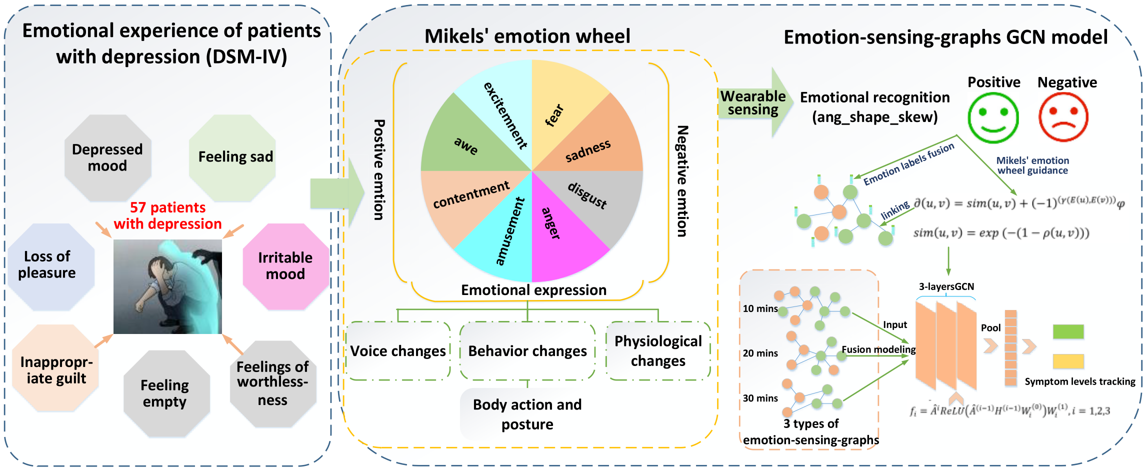


Fig. 4. Overall framework of identifying the levels of changes in depressive symptoms.

and fear, sadness, disgust, and anger belonged to negative emotions are distributed in the right semicircle. Moreover, an individual's emotions can be expressed in their voice, behavior, and physiological changes [26], [27], [28], [29]. Therefore, inspired by this, we have analyzed changes in multimodal features to identify the activity features associated with emotions based on emotion data from healthy controls and thus incorporated subjects' emotions into modeling. The framework of ESG-GCN model is presented in Fig. 4.

1) Emotional Recognition: Generally, patients with depression show low emotion, decreased interest in things, and emotional instability. Emotional recognition is helpful for screening and treating depression. In this study, we have analyzed changes in multimodal features in healthy controls and patients with depression, as well as in patients with depression before and after a course of treatment. For healthy controls, the emotion states of the whole day are recorded in the experiment. In this way, we find that audio features and activity features, namely, audio short-term energy and angular velocity shape skewness ("ang_shape_skew"), have significant changes before and after a course of treatment. These two features of healthy controls and patients also have a consistent trend. Therefore, we can infer that the short-term energy and angular velocity shape skewness have a direct relationship with the patient's emotions. However, the values of audio short-term energy fluctuate in a small range, and it is difficult to distinguish the emotion states. The activity feature for "ang_shape_skew" is determined to mark the emotional state of an emotional unit of the subjects. Human emotions are very complicated, so it is difficult to accurately identify the fine-grained emotional states of patients by a feature. Moreover, according to Mikels' emotion wheel, there are positive and negative polarities in emotions. Positive emotions are distributed in the left semicircle, while negative emotions are distributed in the right semicircle. Therefore, in our research, emotions are divided into two categories, namely, positive and negative emotions.

2) Emotion-Sensing-Graphs Construction: In this study, eight audio features, 43 activity features, and eight heart rate features are extracted. Although high-dimensional features implicitly provide more information, there are also many redundant features. In addition, the feature dimension is too high, while unnecessary noise interference can be introduced so that the performance of the model is worsened. Therefore, we apply PCA to fuse multimodal features and select low-dimensional features of audio, activity, and heartrate that are correlated to depression. Since the patient's emotions are constantly changing, the data are divided into intervals of every 10, 20, or 30 min. We then determine the emotional states for each interval. This approach allows us to build emotion-sensing-graphs based on the relationship between the changes and evolution of emotions and depression.

According to the characteristics of the experimental data, we define an undirected emotion-sensing-graph $G_{IG} := (V, E, H)$, where V is the set of nodes, namely, an emotional unit is regarded as a node, E is the set of edges between nodes, and H is the features of the nodes. The network topology of graph G_{IG} is defined by the adjacency matrix $A \in R^{|V|\times|V|}$ where $A_{u,v} = 1$ if $e_{uv} \in E$ else $A_{u,v} = 0$. Currently, three types of measures, such as Pearson correlation, k-nearest neighbor (KNN), and distance-based rules, are applied to represent the connection relationships of graphs. In this work, the relationship among emotions is measured by Pearson correlation methods. We apply an exponential function to scale the elements. According to Mikels' emotion wheel, adjacent emotions are similar to each other, while opposite emotions are opposite to each other. We have added an emotional guidance item where the same type of emotions will increase the connection weight of emotion-sensing-graphs. In order to obtain a clearer indication of an individual's emotional of changes and connection structure throughout the day, thresholds are applied to select the connection weights. These are based on the following expressions:

$$\partial (u, v) = \mathrm{sim}(u, v) + (-1)^{(\gamma(E(u), E(v)))} \varphi \tag{10}$$

$$\mathrm{sim}(u, v) = \exp(-(1 - \rho(u, v))) \tag{11}$$

$$A_{u,v} = \begin{cases} 1, & \text{if } \partial (u, v) \geq k \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

where $\rho(u, v)$ is the correlation coefficients between two emotional nodes u and v, the term γ is the Kronecker delta function, $E(u)$ and $E(v)$ are the emotional states of

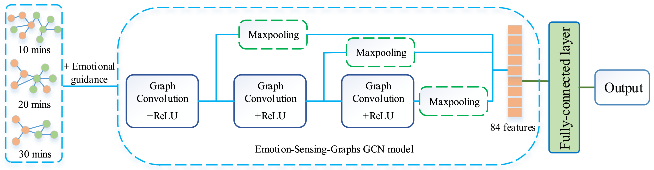


Fig. 5. Architecture of ESG-GCN model.

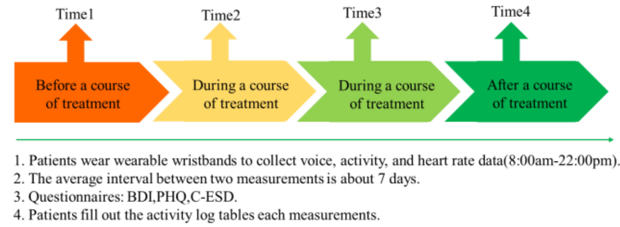


Fig. 6. Experimental process for hospitalized patients with depression.

emotional units, and φ is the emotional similarity coefficient. In particular, the threshold k should satisfy a moderate positive correlation among emotion nodes and maintain a moderate connection of emotion-sensing-graphs.

3) Emotion-Sensing-Graphs Learning: Depression can often reoccur after treatment, and it is difficult for doctors to monitor the patient's state over time. Therefore, monitoring and assessment for the status of patients through wearable sensing wristbands facilitates their treatment and prevents further deterioration. In this article, we have developed emotion-sensing-graphs GCN model (ESG-GCN) to assess the patient's depression levels. The basic framework of the learning model is shown in Fig. 5. One of the key components for a day's emotions is how to learn the relationship and the change of emotions. By now, we have established emotion-sensing-graphs to describe the relationship among emotions. Therefore, the structures among emotional nodes and the features of emotional nodes can be analyzed by spectral graph convolution [30]. Spectral graph convolution decomposes the graph signal $x \in R^N$ in the spectral domain and then applies a spectral filter g_θ on the spectral component to define convolution

$$g_\theta x \approx \sum_{k=0}^{K} \theta'_k T_k \left(L_{\mathrm{sym}}\right) x \tag{13}$$

where $x \in R^N$ is the signal on graph, g_θ is a spectral filter, $\odot$ represents the convolution operation, T_k is the Chebyshev polynomials, θ'_k is Chebyshev coefficients, and L_{sym} is the symmetric Laplacian.

Given an emotion-sensing-graph G, we can calculate its adjacency matrix A and degree matrix $D \in R^{N\times N}$. Kipf and Welling [31] improved the above formula and limited K to 1, which simplifies to

$$g_\theta x \approx \theta \left(I + L_{\mathrm{sym}}\right) x = \theta L x \tag{14}$$

where I represents an identity matrix. In order to enhance the numerical stability of model training, L is normalized

$$H^{(l+1)} = \delta \left(D^{-1/2} L D^{-1/2} H^{(l)} W^{(l)}\right) \tag{15}$$

where $H^{(l)} \in R^{N\times F}$ is the matrix of activation in the lth layer, $H^{(0)}$ is the node input features, and δ is an activation function. In this study, we select the rectified linear unit (ReLU) function. In order to fully mine the information of emotion-sensing-graphs, multilayer GCN networks have been applied to build learning models. The training model consists of three GCN layers

$$f_i = \hat{A}^i \mathrm{ReLU} \left(\hat{A}^{(i-1)} H^{(i-1)} W_i^{(0)}\right) W_i^{(1)}, \quad i = 1, 2, 3 \tag{16}$$

where $\hat{A} = D^{-1/2} L D^{-1/2}$ is a symmetric adjacency matrix, $W_i^{(0)}$ is the weight matrix from the input layer to the hidden layer, and $W_i^{(1)}$ is the weight matrix from the hidden layer to the output layer. Through the three-layer GCN models, the key information of emotion-sensing-graph has been learned. However, in order to obtain more information related to depression, we apply the max-pooling operation on the updated emotion-sensing-graphs learned from each layer of ESG-GCN model and then concatenate the pooled features. Finally, the fully connected layer is employed to identify the depression levels.

The quality of a model is usually evaluated by precision, recall rate, and F1-score where the calculation formulas are defined as follows:

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{17}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{18}$$

$$F1 - \mathrm{score} = 2\frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}} \tag{19}$$

where TP is true positive, FP is false positive, and FN is false negative.

D. Study

1) Subjects: In this study, we have recruited 57 hospitalized patients with depression, with an average age of 16.4 $\pm$ 3.29 years, from the Fourth People's Hospital of Chengdu in China. In addition, 21 healthy individuals with an average age of 19.8 $\pm$ 1.89 years have been selected from a local university to serve as a control group. Patients are diagnosed by an experienced psychiatrist according to DSM-V criteria. All subjects have provided informed consent and signed the agreement.

2) Support Assessment: Multiple questionnaires are used to evaluate the severity of depressive symptoms in both patients with depression and healthy controls. The HAMD rating scale [32] is utilized to assess the severity of depressive symptoms in patients. The patient health questionnaire (PHQ) [33] serves as a straightforward and effective means of screening for depression. It consists of a total of nine items and participants selected the answer that best reflected their feelings in the past two weeks. The beck depression inventory (BDI) questionnaire [34] is employed to determine whether participants are depressed and the severity of their depression. Finally, the Center for Epidemiological Studies Depression (CES-D) Scale [35] is specifically designed to assess the frequency of current depressive symptoms, with an emphasis on depressive mood or affect. The use of multiple questionnaires enables a more comprehensive assessment of depressive symptoms in the study.

TABLE II
DESCRIPTIVE ANALYSIS (MEAN ± SD)

Variables	Time1	Time2	Time3	Time4	Healthy
Age	16.4±3.29				19.8±1.89
Interval (days)	/	7.83±2.48	14.63±3.71	19.97±6.95	
HAMD	26.16±6.15				
PHQ	20.06±5.91	16.21±7.41	16.26±6.88	13.28±8.05	3.67±2.24
BDI	40.76±13.85	35.75±16.64	37.36±17.49	27.72±18.71	6.33±5.18
CES-D	42.04±11.89	36.45±13.72	34.11±13.38	30.43±13.74	11.38±7.26

3) Study Setup: Subjects in this study include hospitalized patients with depression and healthy controls. Detailed study procedures and equipment usage are explained to them, and all procedures are conducted by professional psychologists. Prior to the study, subjects are given a preliminary interview by the researchers to obtain an initial evaluation. The experiments are approved by the Ethical Committee of the University of Electronic Science and Technology of China. The experimental steps of hospitalized patients with depression are shown in Fig. 6. At the beginning of the study, each subject comes to a quiet room, and a professional guides them the test details and demonstrates how to wear the wristband. Next, the subject wears the wristband and remains silent for 2 min in order to collect baseline data. During hospitalization, patients with depression wear the wristbands for three or more days to record data. Since there is little change in the status of healthy controls, we only need to collect data two or three times. For the healthy controls, some variables are controlled such as no extensive activity and no excessive strenuous exercise for the healthy subjects on the day of collection. These factors ensure to some extent that data are collected from patients with depression and healthy control groups in a basically consistent environment. In addition, their emotional states throughout the day are recorded for the healthy controls. On each individual day, the subjects correctly wear the wristband from 8:00 A.M. to 10:00 P.M. and complete an activity log table. The interval between two sessions is about a week. Correspondingly, a series of questionnaires are filled out to record changes in subjects' psychological state. At each measurement, subjects are asked to complete questionnaires, including PHQ, BDI, and CES-D.

III. EVALUATION AND ANALYSIS

A. Descriptive Analysis

The description of age, average inpatient days, and clinical characteristics is shown in Table II. For Time 1, all the 57 subjects are assessed. For Times 2 and 3, 24 and 19 patients, respectively, have completed the questionnaire during hospitalization. For Time 4, due to the impact of COVID-19, only 39 patients are assessed. In addition, 21 healthy subjects have participated in our control trials. Paired sample t-tests show significant differences between before and after the course of treatment during hospitalization on the scores of PHQ ($t = 6.4$, $p < 0.001$), BDI ($t = 4.9$, $p < 0.001$), and CES-D ($t = 5.9$, $p < 0.001$). It can be seen from Table I that after a course of treatment, the PHQ scores of patients with depression decrease, which also indicates that the depressive symptoms of patients are improved to a certain extent.

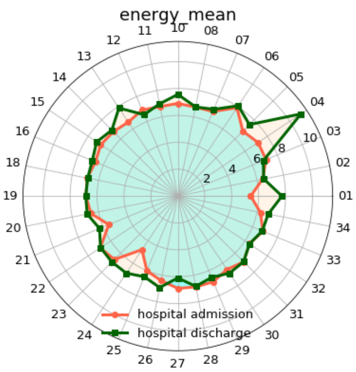


Fig. 7. Audio short-time energy for hospital admission and discharge of patients with depression.

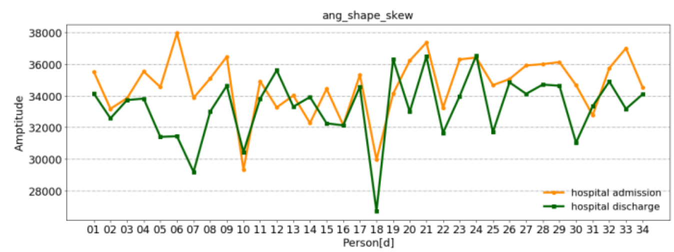


Fig. 8. "ang_shape_skew" for hospital admission and discharge of patients with depression.

B. Emotional Recognition Analysis

1) Comparison of One-Day Data for Patients With Depression: In Figs. 7 and 8, the features of audio short-time energy and "ang_shape_skew" for the hospital admission and discharge of patients with depression are compared. In order to clearly investigate the changes of multimodal features, five patients whose PHQ scores do not improve after a course of treatment are not shown in the figure. Due to small technical error, the audio data of two patients are not captured, so only the audio features of 32 patients with depression are analyzed. We can clearly see from Fig. 6 that audio short-time energy ($p = 0.005$) at hospital discharge for depressed patients is generally high, and its proportion reaches 81%. The audio short-time energy represents the energy of the audio signal, and it can easily distinguish whether the patients have sound signals. In addition, we can conclude that the "ang_shape_skew" ($p = 0.0002$) for hospital discharge of patients is almost reduced and its proportion reaches 82%. The feature represents the shape statistical features of PSD for angular velocity. This might indicate that patients with depression are more emotionally stable and calm after a course of treatment. Therefore, the features of audio short-time energy and "ang_shape_skew" have significant changes with the remission of depressive symptoms. This also shows that multimodal wearable sensing devices we developed are effective in tracking and monitoring the changes of depression symptoms.

2) Analysis of Emotional Changes for Patients With Depression: In Fig. 9, the mean of "ang_shape_skew" feature for all

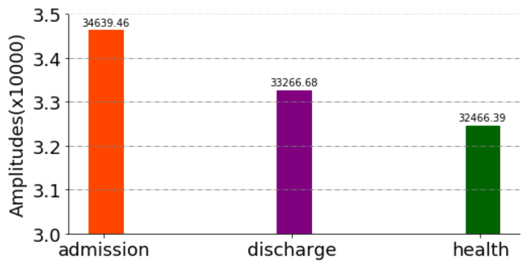


Fig. 9. Comparisons of "ang_shape_skew" for hospital admission and discharge of patients with depression and healthy controls.

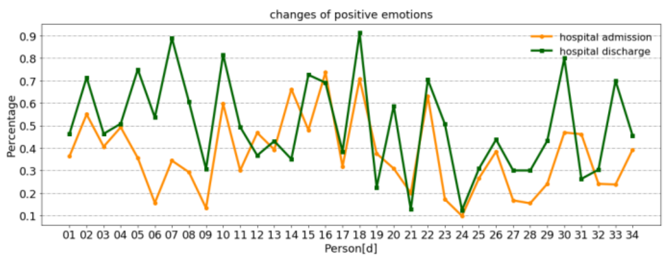


Fig. 10. Changes of positive emotions for hospital admission and discharge of patients with depression.

patients before and after a course of treatment and healthy controls is compared. As can be seen from the figure, when patients are first hospitalized, the mean of "ang_shape_skew" feature is higher. After a course of treatment, the mean of "ang_shape_skew" features for patients presents a decrease. Moreover, we can also clearly find that the average value of "ang_shape_skew" is at a lower level in healthy controls. Thus, from the statistical significance, we can conclude that the mean of "ang_shape_skew" is lower when positive emotions are the main counterpart, while the mean of "ang_shape_skew" is higher when patients' negative emotions increase. In Fig. 10, the 10-min data are defined as an emotional unit and we have analyzed the emotional changes of patients with depression throughout the day before and after a course of treatment. As shown in the figure, the proportion of positive emotions throughout the day shows an upward trend and the proportion of patients reaches 82%. This also indicates that the negative emotions of patients with depression gradually decrease with the relief of depressive symptoms.

C. Analysis of Emotion-Sensing-Graphs

In Fig. 11, emotion-sensing-graphs of hospital admission and discharge for patients with depression and health controls are compared. As can be seen in Fig. 10, the same emotions are more closely linked to each other. Concurrently, we can find that the numbers of positive emotion nodes for patients with depression after a course of treatment and healthy controls have increased in the emotion-sensing-graphs. The results also show that negative emotions for patients with depression occupy more time in a day before a course of treatment. This may be associated with patients with depression being in a poor mood or often in a quiet state. Table III also indicates that the average degree ($t = -2.41$ and $p = 0.023$) and proportion ($t = -2.82$ and $p = 0.009$) of positive emotion nodes for patients with depression before and after a course of treatment have increased significantly. Moreover, the healthy controls

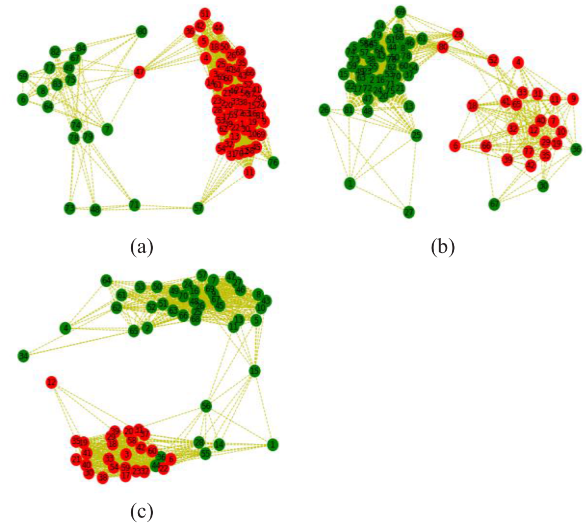


Fig. 11. Emotion-sensing-graphs. (a) Hospital admission for patients with depression. (b) Hospital discharge for patients with depression. (c) Healthy control. (Red nodes represent negative emotions; green nodes represent positive emotions.)

TABLE III
POSITIVE EMOTION CHANGES FOR EMOTION-SENSING-GRAPHS (MEAN ± SD)

	Time1	Time4	Healthy
Degrees of positive emotion nodes	7.65±4.77	10.61±5.79*	9.87±5.08
Proportion of positive emotion nodes	0.37±0.21	0.50±0.22*	0.57±0.22

Note: * represents the p (< 0.05) value of paired sample t-test, Time 1 compare to Time 4 for patients with depression.

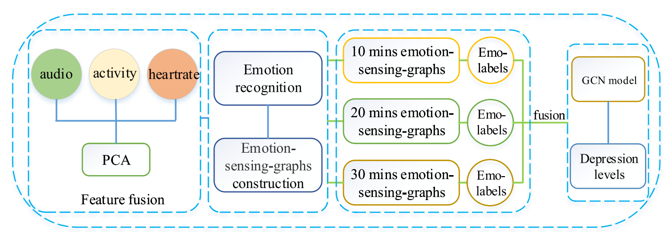


Fig. 12. Framework of depression levels assessment for ESG-GCN model.

have a higher degree of connectedness of positive emotion nodes and positive emotion nodes are the main component in the emotion-sensing-graphs. Therefore, the emotion-sensing-graphs can, to some extent, reflect the levels of depressive symptoms in patients with the relief of depressive symptoms.

D. Evaluation of Depression Symptoms

1) Evaluation Model of Emotion-Sensing-Graphs: The framework of an emotion-sensing-graph learning model is shown in Fig. 12. We average the extracted features for each emotional unit, such as 10, 20, and 30 min, to obtain features that represent the overall trend emotional state of each emotional unit. Then, eight audio features, 43 activity features, and eight heart rate features are spliced to obtain a high-dimensional vector of 59 features. However, these features have considerable redundancy and are unsuitable for use in machine learning models. In addition, the mental state is often manifested by a variety of social signals such as

TABLE IV
DIVIDE PHQ LEVELS

Depression Levels (Label)	PHQ scores
0	15-27
1	0-14

audio, activity, and physiology, and it is therefore interesting to fuse multimodal features for analysis. We apply a PCA model to obtain ten fusion features. Generally, a person's emotional state is directly related to their mental health. Most severely depressed patients are in a chronically negative mood. Therefore, according to Mikels' emotion wheel, we have constructed three types of emotion-sensing-graphs based on emotional states and the relationship between them. During model construction, emotional state labels are integrated into the emotional node attributes of emotion-sensing-graphs and the emotional node dimension of the emotion-sensing-graphs is 11. The GCN model learns both emotional nodes' information and connections among emotional nodes. Thus, the three-layer GCN model is developed to learn three types of emotion-sensing-graphs to identify and evaluate the depression symptom levels in patients.

2) Evaluation of Depression Severity: Subjects filled out the PHQ questionnaire during this study, and we can judge the levels of depression based on their scores. During a course of treatment, the classification of depression levels can often be ambiguous, particularly when patients experience significant deviations between moderately severe and severe depression or when subjective judgments made by patients with depression are imprecise. Furthermore, we focus on the changes in symptoms of patients with depression, namely, the improvement of depression during the treatment process. Considering the advice provided by psychiatrists and the inherent challenges of swiftly transitioning from MDD to a state of no depression, we propose a simplified classification approach. Depression levels are categorized into two classes: Scores of 0–14 represent no depression or mild depression, 15–27 as clinical MDD. In Table IV, we divide depression levels into these two classes.

In this study, considering the synchronization and integrity of audio, activity, and heart rate data, we have, respectively, obtained 153 emotion-sensing-graphs with emotional units of 10, 20, and 30 min for 57 patients with depression and 21 healthy controls. Therefore, a total of 459 emotion-sensing-graphs are used to establish the model. We then develop the ESG-GCN model to identify the changes in the symptoms of patients with depression. The training model consists of three-GCN layers, where the feature dimension of input variable for the first GCN layer is 11, the dimension of middle layer is 32, and the dimension of the third GCN layer is determined as 20. In addition, the learning rate is $2e^{-4}$ and the number of iterations is 1400. To avoid overfitting and improve the performance of the training model, tenfold cross validation is adopted. The dataset is randomly divided into ten parts, with nine parts of them taking turns as training data and one part as testing data for experimentation. Therefore, the training and testing datasets, respectively, are 413 and 46 emotion-sensing-graphs at each experiment. The performance of models for each experiment is presented in Table V. It can be seen that

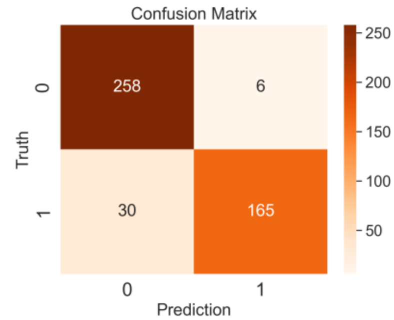


Fig. 13. Confusion matrix for classification of depression levels.

except for two experiments, the precision, recall, F1-scores, and accuracy of ESG-GCN model on the test set exceed 0.8 for each other experiment. Furthermore, the standard deviation of the tenfold cross validation is 0.05. The performance indicators of the model do not fluctuate significantly in each experiment. This indicates that the model has a certain degree of effectiveness and stability.

In Fig. 13, the confusion matrix of the emotion-sensing-graphs learning model is drawn. It can be seen from the classification results that the learning model can effectively identify the changes in symptoms of patients with depression. Concurrently, we can conclude that there is a corresponding relationship between depression and the multimodal data of audio, activity, and heart rate. Hence, this shows that we can use multisensing wearable wristbands to objectively assess the depression levels of patients.

In order to build a reliable prediction model of depression severity, six machine learning models, which are XGBoost RandomForest, DecisionTree, LogisticRegression, KNeighbors, and Naive Bayes (Gaussian), are trained. The performance parameters of six models can be seen in Table VI.

From Table VI, we can find that the precision (0.83 ± 0.05), recall rate (0.82 ± 0.05), F1-scores (0.82 ± 0.05), and validation accuracy (0.83 ± 0.05) of ESG-GCN model are the best, and the stability of the learning model is also more prominent. However, the precision and accuracy of XGBoost, RandomForest, DecisionTree, Naive Bayes (Gaussian), and LogisticRegression models are about 0.6. The KNeighbors model is not suitable for the depression data in this study. Therefore, emotion-sensing-graphs containing the changes of nodes (emotions) and the mutual influence relationships among nodes (emotions) are helpful to improve the model performances for depression data collected naturally over a long period of time. Simultaneously, these results also show that the ESG-GCN model is effective to track the change of depression levels.

In Table VII, the five emotion-sensing-graph schemes are applied to the ESG-GCN model to compare the performances. As seen from the table, the performances of the models built with emotion-sensing-graphs for 10, 20, or 30 min as emotional units are not outstanding. In addition, the tabular results show that the fusion training of three types of emotion-sensing-graphs can significantly improve the performance of the model. Concurrently, we can also find that the precision (0.83 ± 0.05), recall rate (0.82 ± 0.05), F1-scores (0.82 ± 0.05), and validation accuracy (0.83 ± 0.05) of the model

TABLE V
PERFORMANCE OF THE ESG-GCN MODEL FOR TENFOLD CROSS VALIDATION

performance	Each result for 10-fold cross-validation										Mean	Std
	1	2	3	4	5	6	7	8	9	10		
Precision	0.893	0.863	0.815	0.818	0.861	0.739	0.868	0.852	0.806	0.755	0.83	0.05
Recall	0.893	0.842	0.815	0.825	0.861	0.726	0.823	0.869	0.806	0.753	0.82	0.05
F1-scores	0.889	0.842	0.815	0.823	0.861	0.727	0.833	0.858	0.806	0.754	0.82	0.05
Accuracy	0.889	0.844	0.822	0.822	0.867	0.733	0.844	0.867	0.822	0.756	0.83	0.05

TABLE VI
PERFORMANCE OF DIFFERENT MODELS

Models	Precision	Recall	F1-scores	Accuracy
XGBoost	0.67(±0.08)	0.66(±0.07)	0.65(±0.07)	0.66(±0.07)
RandomForest	0.59(±0.13)	0.64(±0.12)	0.59(±0.13)	0.63(±0.1)
DecisionTree	0.68(±0.15)	0.64(±0.13)	0.61(±0.13)	0.64(±0.13)
KNeighbors	0.55(±0.14)	0.56(±0.13)	0.54(±0.14)	0.56(±0.13)
Naive Bayes (Gaussian)	0.68(±0.15)	0.64(±0.13)	0.61(±0.13)	0.64(±0.13)
LogisticRegression	0.71(±0.13)	0.68(±0.15)	0.65(±0.17)	0.68(±0.15)
Proposed model	**0.83(±0.05)**	**0.82(±0.05)**	**0.82(±0.05)**	**0.83(±0.05)**

TABLE VII
PERFORMANCE OF MODELS FOR DIFFERENT EMOTION-SENSING-GRAPHS

Emotion-sensing-graphs	Precision	Recall	F1-scores	Accuracy
10-mins	0.74(±0.1)	0.66(±0.1)	0.61(±0.12)	0.65(±0.11)
20-mins	0.64(±0.14)	0.61(±0.13)	0.59(±0.12)	0.61(±0.11)
30-mins	0.61(±0.16)	0.60(±0.13)	0.58(±0.14)	0.61(±0.14)
Fusion graphs	0.78(±0.05)	0.77(±0.05)	0.77(±0.05)	0.78(±0.05)
Emotional guidance	0.83(±0.05)	0.82(±0.05)	0.82(±0.05)	0.83(±0.05)

established by three types of emotion-sensing-graph with emotional guidance are the highest. Therefore, the results show that emotion-sensing-graphs with emotional guidance play an important role in the monitoring of change in patients' condition during hospitalization. This also indicates the effectiveness of detecting changes in depressive symptoms from the perspective of emotional status.

E. Discussion

In recent years, research on depression recognition based on audio signals has received widespread attention. By analyzing various audio features, the emotional and mental health status of individuals can be objectively evaluated, providing a new approach for early identification and intervention of depression. Some studies have found that the speech of patients with depression often exhibits characteristics of low energy, lack of expressiveness, and negative emotions. In this study, we have found that the short-term energy of audio is an upward trend after a course of treatment. This is also consistent with the findings of some previous studies. For example, earlier studies have also found a correlation between levels of depression and changes in audio features, such as reduced pitch range, loudness, and energy dynamics [36]. In our research, we have also found that the activity features of "ang_shape_skew" show a significant downward trend with the remission of depressive symptoms. It may also be related to the reduction of abnormal mood for patients with depression after a course of treatment. For example, some research has shown that adolescents with depression had significant fluctuations in behavioral expression, such as emotional irritability, irritability, impulsivity, and disobedience to teachers and parents [37], [38], [39]. Therefore, features of audio and activity may become important indicators for analyzing and observing changes in depressive symptoms and this also indicates that we can track the treatment process of patients with depression by monitoring changes in audio and activity data collected by wearable wristbands.

In this study, we construct individual's emotion-sensing-graphs based on emotions. From the results, it can be seen that there are significant differences in emotion-sensing-graphs not only among patients before and after a course of treatment but also between the healthy control groups and patients. This also indicates that emotions can serve as important reference indicators for distinguishing depression and for tracking the changes in symptoms of patients with depression before and after a course of treatment. Usually, an individual's psychological states are easily detected based on their emotional states in real life. For example, long-term negative emotions, such as sadness, pain, inferiority, and worldliness, are largely related to psychological disorders such as depression. Furthermore, it is well known that depression is a common emotional disorder mental illness, and clinically, prolonged low mood and unhappiness in real life can be observed. In addition, it can be seen from the model training results that the recognition model of changes in depressive symptoms constructed by emotion-sensing-graphs with emotional guidance and emotional label integration has a better performance compared to several machine learning models in this article. Compared to machine learning that only relies on feature attributes, emotion-sensing-graphs that contain changes of emotional state, emotional nodes at different time periods, and relationships among emotional nodes have more comprehensive depression information, which may play an important role in improving the performance of the model. Therefore, analyzing and identifying depression in the form of graph structure based on prior knowledge of psychology and emotional models may be a new approach for wearable technology to assist in the diagnosis and treatment of depression.

F. Conclusion

In this study, we have developed a wearable sensing wristband to dynamically monitor and track the treatment process of patients with depression. The wearable sensing wristbands integrate audio, accelerometer and gyroscope, and heart rate sensors. Using this system, we have conducted a comparative experiment between hospitalized patients with depression and healthy controls and developed feature extraction and comparative analysis on the multimodal data. We find that there is a significant increase in audio short-term energy, while the activity feature of "ang_shape_skew" shows a decreasing trend after a course of treatment. A person's emotions are directly related to their behavior and mental state. We define subjects' daily data as three emotional units and identify whether each emotional unit's emotional state is positive or negative. Subsequently, in order to further explore the evolution of emotions and the relationship between emotions, emotion-sensing-graphs guided by Mikels' emotion wheel are established using the correlation among emotions. The results show that the same emotions are more closely connected and the average degree and proportion of positive emotion nodes after a course of treatment have increased significantly in emotion-sensing-graphs. To further identify change in each patient's condition, we have developed the ESG-GCN model. The obtained results have indicated that the accuracy and stability of our designed model have good levels of performance for the assessment of depression severity. It is also found that the fusion training of three types of emotion-sensing-graphs with emotion labels can further improve the performance of the model. This also indicates that emotions play an important role in the recognition of depressive symptoms. In future research, detailed experiments will be performed on recognition, assessment, and monitoring of depression to improve the reliability of the proposed system.

References

[1] (2020). *World Health Organization*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[2] E. Saito and A. Van Meter, "25.2 comparison of depressive and manic symptoms among youth with depression, bipolar disorder, and those who convert from depression to bipolar," *J. Amer. Acad. Child Adolescent Psychiatry*, vol. 59, no. 10, p. 304, Oct. 2020.

[3] J. Chen, "Diagnosis, mechanism, and current treatment of depression: Current situation and future directions," *Int. J. Pharma Med. Biol. Sci.*, vol. 10, no. 3, pp. 125–129, 2021.

[4] F. Hardeveld, J. Spijker, R. De Graaf, W. A. Nolen, and A. T. F. Beekman, "Prevalence and predictors of recurrence of major depressive disorder in the adult population," *Acta Psychiatrica Scandinavica*, vol. 122, no. 3, pp. 184–191, Sep. 2010.

[5] K. P. Ebmeier, C. Donaghey, and J. D. Steele, "Recent developments and current controversies in depression," *Lancet*, vol. 367, no. 9505, pp. 153–167, Jan. 2006.

[6] D. J. France, R. G. Shiavi, S. Silverman, M. Silverman, and M. Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 829–837, Jul. 2000.

[7] J. Jin, B. Gao, S. Yang, B. Zhao, L. Luo, and W. L. Woo, "Attention-block deep learning based features fusion in wearable social sensor for mental wellbeing evaluations," *IEEE Access*, vol. 8, pp. 89258–89268, 2020.

[8] K. Ueafuea et al., "Potential applications of mobile and wearable devices for psychological support during the COVID-19 pandemic: A review," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7162–7178, Mar. 2021.

[9] J.-Y. Kim, N. Liu, H.-X. Tan, and C.-H. Chu, "Unobtrusive monitoring to detect depression for elderly with chronic illnesses," *IEEE Sensors J.*, vol. 17, no. 17, pp. 5694–5704, Sep. 2017.

[10] S. Kontaxis et al., "Photoplethysmographic waveform analysis for autonomic reactivity assessment in depression," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 4, pp. 1273–1281, Apr. 2021.

[11] Z. Wang, L. Chen, L. Wang, and G. Diao, "Recognition of audio depression based on convolutional neural network and generative antagonism network model," *IEEE Access*, vol. 8, pp. 101181–101191, 2020.

[12] 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.

[13] N. R. Leonard et al., "Mobile health technology using a wearable sensorband for female college students with problem drinking: An acceptability and feasibility study," *JMIR mHealth uHealth*, vol. 5, no. 7, pp. 79–90, Jul. 2017.

[14] A. Lanata, G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo, "Complexity index from a personalized wearable monitoring system for assessing remission in mental health," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 132–139, Jan. 2015.

[15] W. Wu, H. Zhang, S. Pirbhulal, S. C. Mukhopadhyay, and Y.-T. Zhang, "Assessment of biofeedback training for emotion management through wearable textile physiological monitoring system," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7087–7095, Dec. 2015.

[16] Y. Tazawa et al., "Evaluating depression with multimodal wristband-type wearable device: Screening and assessing patient severity utilizing machine-learning," *Heliyon*, vol. 6, no. 2, Feb. 2020, Art. no. e03274.

[17] P. Pedrelli, S. Fedor, A. Ghandeharioun, E. Howe, and R. W. Picard, "Monitoring changes in depression severity using wearable and mobile sensors," *Front. Psychiatry*, vol. 11, Dec. 2020, Art. no. 584711.

[18] R. Wang et al., "Tracking depression dynamics in college students using mobile phone and wearable sensing," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–26, Mar. 2018.

[19] P. Bidja, "DepressionGNN: Depression prediction using graph neural network on smartphone and wearable sensors," Honors Scholar Theses, Tech. Rep. 607, 2019. [Online]. Available: https://digitalcommons.lib.uconn.edu/srhonors_theses/607

[20] T. Chen, Y. Guo, S. Hao, and R. Hong, "Exploring self-attention graph pooling with EEG-based topological structure and soft label for depression detection," *IEEE Trans. Affect. Comput.*, vol. 13, no. 4, pp. 2106–2118, Oct. 2022.

[21] Z. Tariq, S. K. Shah, and Y. Lee, "Speech emotion detection using IoT based deep learning for health care," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2019, pp. 4191–4196.

[22] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, and A. Reuter-Lorenz, "Emotional category data on images from the international affective picture system," *Behav. Res. Methods*, vol. 37, no. 4, pp. 626–630, 2005.

[23] J. Gu et al., "Wearable social sensing: Content-based processing methodology and implementation," *IEEE Sensors J.*, vol. 17, no. 21, pp. 7167–7176, Nov. 2017.

[24] L. Jiang et al., "Wearable long-term social sensing for mental wellbeing," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8532–8542, Oct. 2019.

[25] *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV*. Washington, DC, USA: American Psychiatric Association, 1994.

[26] M. de Meijer, "The contribution of general features of body movement to the attribution of emotions," *J. Nonverbal Behav.*, vol. 13, no. 4, pp. 247–268, Dec. 1989.

[27] Z. Jiang, S. Harati, A. Crowell, H. S. Mayberg, S. Nemati, and G. D. Clifford, "Classifying major depressive disorder and response to deep brain stimulation over time by analyzing facial expressions," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 2, pp. 664–672, Feb. 2021.

[28] M. Sawada, K. Suda, and M. Ishii, "Expression of emotions in dance: Relation between arm movement characteristics and emotion," *Perceptual Motor Skills*, vol. 97, no. 3, pp. 697–708, Dec. 2003.

[29] C. L. Roether, L. Omlor, A. Christensen, and M. A. Giese, "Critical features for the perception of emotion from gait," *J. Vis.*, vol. 9, no. 6, p. 15, Jun. 2009.

[30] F. R. K. Chung, *Spectral Graph Theory*, vol. 92. Providence, RI, USA: American Mathematical Society, 1997.

[31] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent.*, 2017, pp. 1–14.

[32] M. Hamilton, "A rating scale for depression," *J. Neurol. Neurosurgery*, vol. 23, no. 1, pp. 56–62, 1960.

[33] K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The PHQ-9: Validity of a brief depression severity measure," *J. Gen. Internal Med.*, vol. 16, no. 9, pp. 606–613, Sep. 2001.

[34] A. T. Beck, R. A. Steer, and M. G. Carbin, "Psychometric properties of the beck depression inventory: Twenty-five years of evaluation," *Clin. Psychol. Rev.*, vol. 8, no. 1, pp. 77–100, Jan. 1988.

[35] L. S. Radloff, "The CES-D scale: A self-report depression scale for research in the general population," *Appl. Psychol. Meas.*, vol. 1, no. 3, pp. 385–401, Jun. 1977.

[36] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. F. Quatieri, "A review of depression and suicide risk assessment using speech analysis," *Speech Commun.*, vol. 71, pp. 10–49, Jul. 2015.

[37] A. R. I. Solomon, D. A. F. Haaga, and B. A. Arnow, "Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research," *J. Nervous Mental Disease*, vol. 189, no. 8, pp. 498–506, 2001.

[38] M. Crowe, N. Ward, B. Dunnachie, and M. Roberts, "Characteristics of adolescent depression," *Int. J. Mental Health Nursing*, vol. 15, no. 1, pp. 10–18, 2006.

[39] B. Nardi, G. Francesconi, M. Catena-Dell'osso, and C. Bellantuono, "Adolescent depression: Clinical features and therapeutic strategies," *Eur. Rev. Med. Pharmacolog. Sci.*, vol. 17, no. 11, pp. 1546–1551, Jun. 2013.

Daili Yang received the B.Sc. degree from the School of Physics and Electronic Engineering, Hainan Normal University, Haikou, China, in 2009, and the M.Sc. degree from the School of Electrical Engineering, Xinjiang University, Urumqi, China, in 2014. He is currently pursuing the Ph.D. degree in control science and engineering with the University of Electronic Science and Technology of China, Chengdu, China.

His research interests include wearable sensors and physical and mental health.

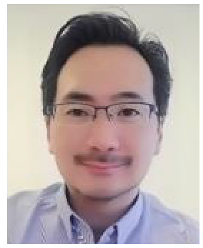
Yunge Li received the B.Sc. degree from the Department of Clinical Medicine, West China Medical University, Chengdu, China, in 1996, and the M.Sc. degree from the West China Clinical Medical College, Sichuan University, Chengdu, in 2016.

She is an Associate Professor with the Clinical Hospital of Chengdu Brain Science Institute, University of Electronic Science and Technology of China (UESTC), Chengdu, and the Deputy Chief Physician of the Clinical Psychology Department, Fourth People's Hospital, Chengdu. Her research interests include clinical psychology.

Bin Gao (Senior Member, IEEE) received the B.Sc. degree in communications and signal processing from Southwest Jiao Tong University, Chengdu, China, in 2005, and the M.Sc. degree (Distinction) in communications and signal processing and the Ph.D. degree from Newcastle University, Newcastle upon Tyne, U.K., in 2006 and 2011, respectively.

From 2011 to 2013, he worked as a Research Associate of wearable acoustic sensor technology with Newcastle University. He is currently a Professor with the School of Automation Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu. His research interests include sensor signal processing, multiphysical sensing and instrument, social wearable sensing, nondestructive testing, and evaluation where he actively publishes in these areas.

Dr. Gao is also a very active reviewer for many international journals and long-standing conferences. He has coordinated several research projects from the National Natural Science Foundation of China.



Wai Lok Woo (Senior Member, IEEE) received the B.Eng. degree in electrical and electronics engineering and the M.Sc. and Ph.D. degrees in statistical machine learning from Newcastle University, Newcastle upon Tyne, U.K., in 1993, 1995, and 1998, respectively.

He was the Director of Research with the Newcastle Research and Innovation Institute and the Director of Operations with Newcastle University. He is currently a Professor of Machine Learning with Northumbria University, Newcastle upon Tyne. He has published more than 400 articles on these topics on various journals and international conference proceedings. His research interests include mathematical theory and algorithms for data science and analytics, artificial intelligence, machine learning, data mining, latent component analysis, multidimensional signal, and image processing.

Dr. Woo is a member of the Institution of Engineering Technology. He was a recipient of the IEE Prize and the British Commonwealth Scholarship. He serves as an Associate Editor for several international signal processing journals, including *IET Signal Processing*, the *Journal of Computers*, and the *Journal of Electrical and Computer Engineering*.

Yiwei Zhang received the B.Sc. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018, where she is currently pursuing the M.Sc. degree in human activity recognition and machine translation.

Her research interests include wearable sensing, recommendation systems, and machine learning.

Keith M. Kendrick received the B.A. and Ph.D. degrees in psychology from the University of Durham, Durham, U.K., in 1976 and 1979, respectively.

He has held research positions with the University of Durham, the Institute of Zoology, London, U.K.; University of Cambridge, Cambridge, U.K.; and the Babraham Institute, Babraham, U.K., prior to moving his current post with the School of Life Science and Technology, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2011, as a National 1000 Talent Professor. He has published more than 240 peer-reviewed papers, including many in Nature, Science, and PNAS as well as in leading Psychiatric journals. The total citation for his publications is now more than 14 500, with 40 articles' citations of more than 100 and five articles of more than 300 and the H index is 71. He is titled as *Elsevier China Highly Cited Scholars* and has received numerous competitive research grants from both U.K.- and China-based funding agencies. His major research is neurotherapy and social cognition and affective neuroscience. This includes finding new and effective detection methods and treatments for mental disorders.

Lizhu Luo received the bachelor's and master's degrees in psychology from Southwest University, Chongqing, China, in 2010 and 2012, respectively, and the Ph.D. degree in biomedical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2017.

She is an Associate Professor with the Clinical Hospital of Chengdu Brain Science Institute, UESTC. She is a Psychologist and a Neuroscientist with a strong interest in affective neuroscience and clinical psychology. She further honed her expertise through Postdoctoral Research with the Clare Hall and the Department of Psychiatry, University of Cambridge, Cambridge, U.K. Her research was funded by the National Natural Science Foundation of China, the China Postdoctoral Science Foundation, and the Sichuan Provincial Science and Technology Foundation.