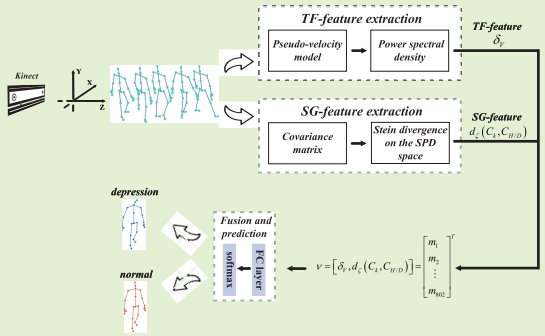
A Gait Assessment Framework for Depression Detection Using Kinect Sensors

Tao Wang, *Graduate Student Member, IEEE*, Cancheng Li, Chunyun Wu, *Graduate Student Member, IEEE*, Chengjian Zhao, *Graduate Student Member, IEEE*, Jieqiong Sun, Hong Peng, Xiping Hu, *Member, IEEE*, and Bin Hu, *Senior Member, IEEE*

***Abstract*—As depression becomes more commonplace in society, the timely and effective detection of the signs of depression for its prevention and early treatment becomes more important. Gait analysis can provide a contactless and low-cost method for depression diagnosis. In this study, we propose a novel gait assessment framework to implement non-intrusive, real-time and automatic depression detection using Kinect, an inexpensive and portable depth sensor. We focus on extracting a novel time-domain and frequency-domain feature (TF-feature) and a spatial geometric feature (SG-feature), and investigating the effectiveness of fused features in detecting depression for the non-contact gait data. A pseudo-velocity model is firstly built to analyze the gait abnormalities of individuals with depression in the time domain. Subsequently, we perform the power spectral density (PSD) analysis on the model to extract the TF-feature. Then, the covariance matrices and the symmetric Stein divergence (S-divergence) are leveraged to obtain the SG-feature, which is fused with TF-feature to form new features for classification. The experimental results on 95 subjects (43 scored-depressed and 52 non-depressed individuals) show that the proposed method achieves a good classification accuracy of 93.75%, has superior performance compared to several other methods, and significantly alleviates the impact of individual differences. These results indicate the efficacy and robustness of the proposed framework for depression detection.**



***Index Terms*—Gait, depression detection, Kinect sensor, pseudo-velocity model, time-frequency feature, covariance matrices.**

Manuscript received August 17, 2020; accepted September 3, 2020. Date of publication September 7, 2020; date of current version January 6, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFA0706200; in part by the National Natural Science Foundation of China under Grant 61632014, Grant 61627808, and Grant 61210010; in part by the National Basic Research Program of China (973 Program), under Grant 2014CB744600; in part by the Program of Beijing Municipal Science and Technology Commission under Grant Z171100000117005; and in part by the Fundamental Research Funds for the Central Universities under Grant lzujbky-2020-kb25. The associate editor coordinating the review of this article and approving it for publication was Prof. Rosario Morello. *(Corresponding authors: Hong Peng; Xiping Hu; Bin Hu.)*

Tao Wang, Cancheng Li, Chunyun Wu, Chengjian Zhao, Jieqiong Sun, and Hong Peng are with the Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China (e-mail: wangtao2018@lzu.edu.cn; licch17@lzu.edu.cn; wuchy18@lzu.edu.cn; zhaochj18@lzu.edu.cn; sunjq18@lzu.edu.cn; pengh@lzu.edu.cn).

Xiping Hu is with the Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China, and also with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shengzhen 518055, China (e-mail: huxp@lzu.edu.cn).

Bin Hu is with the Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China, also with the CAS Center for Excellence in Brain Science and Institutes for Biological Sciences, Shanghai Institutes for Biological Sciences, Chines Academy of Sciences, Shanghai 200031, China, and also with the Beijing Institute for Brain Disorders, Capital Medical University, Beijing 100069, China (e-mail: bh@lzu.edu.cn).

Digital Object Identifier 10.1109/JSEN.2020.3022374

I. Introduction

DEPRESSION is a common disease in psychiatry. According to the World Health Organization, clinical depression has affected approximately 264 million people globally [1]. However, the existing depression treatment studies have mostly focused on rehabilitation and have neglected early detection, which can effectively prevent the occurrence of more serious negative consequences [2]. Therefore, in recent years, interest in automatic depression assessment has grown rapidly.

The human gait is a natural behavior composed of several repetitive parts, known as gait cycles. It provides more useful information than only mobility and has been applied in the analysis of depression. Compared with other methods in current depression detection, such as those based on audio, text, magnetic resonance imaging (MRI) and electroencephalogram (EEG) [3]–[7], methods based on gait have the following advantages. First, gait does not require high-resolution capture or special equipment [8]. Second, gait movement can be captured from a distance, enabling a contactless method; only the simple cooperation of the subject is needed.

Because of these advantages, researchers have proposed some sensor-based gait analysis methods to study depression. In [9], the authors used a video camera and special devices

1558-1748 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

placed on shoes to collect the gait data of subjects and found a linear correlation between walking velocity and depression severity. Utilizing a motion capture system with video cameras and markers attached to the participants' bodies, some gait abnormalities of depressed individuals were found in [10], such as reduced walking velocity, increased body sway and more slumped posture, compared to healthy individuals. With a digital camera, the study reported in [11] also verified that slouching postures were a prominent feature of patients with depression during gait movements.

With the advent of low-cost and non-intrusive motion caption sensors, such as Microsoft Kinect, we can trace the major joints (Kinect V2 provides 25 points) of the human body dynamics accurately in a 3D (x-, y- and z-axes) manner without the assistance of additional equipment or the requirement of a specifically designed environment [12]. Compared with the traditional gait acquisition equipment, such as human perceivers (e.g., inertial measurement unit and smart markers) and RGB cameras, the gait data collected by Kinect are recorded in a non-contact manner, which can avoid the potential effects of markers and wearable devices. In addition, the depth maps collected by Kinect are insensitive to illumination changes and contain more 3D information [13]. Hence, they have proven to be reliable in estimating body skeletal information [14]. As an ideal candidate for a low-cost system, the Kinect device can be effectively employed for action recognition [15], affective computing [16], gait analysis [17], and home rehabilitation and monitoring [18].

The recently introduced Kinect-captured gait data have been widely implemented in clinical applications, such as those involving elderly people at home who have a high risk of falling, individuals with multiple sclerosis and patients suffering from Parkinson's disease [19]–[21]. Furthermore, Kinect has been used in gait-based depression detection. Leveraging the fast Fourier transform (FFT) to extract features from gait data collected by Kinect, the authors of [22] trained regression models to recognize the level of anxiety and depression. In [23], the authors used Kinect to capture the gait trajectory and proposed a new direction for detecting depression with gait frequency features based on the Hilbert-Huang transform. Reference [24] also used Kinect to collect gait skeletal data and investigated the associations between gait abnormalities and depression.

Although some studies have utilized 3D gait skeleton data extracted by Kinect for depression detection, some issues persist with the current data analysis. First, the existing method for extracting the gait abnormalities of patients with depression concentrates on a certain dimensional coordinate of a few special joints and only focuses on the frames with extremely abnormal gait. This approach renders the consecutive 25-joint frame sequence collected by Kinect to be only partially utilized, with abnormalities in several gait characteristics ineffectively expressed. Thus, a particular model to analyze gait anomalies is urgently needed. Furthermore, the current feature extraction methods are analyzed in the time domain and the frequency domain separately. Features in both domains need to be considered comprehensively to analyze the gait data from patients with depression. Second, in the current gait research based on Kinect for depression detection, each joint is studied independently, ignoring the relationships between joints. These relationships might enable extracting the irregular and complex gait spatial information for skeleton-based recognition tasks [25].

To address these issues, in this paper, a gait assessment framework implementing non-intrusive, real-time and automatic depression detection is proposed. The input of our system is composed of 3D skeleton sequences provided by two Kinect V2 devices attached to the ceiling. First, we conduct preprocessing to improve the data quality. Subsequently, time-domain and frequency-domain feature (TF-feature) and spatial geometric feature (SG-feature) extractions are performed on the data. In the extraction of TF-feature, we construct a pseudo-velocity model to analyze the gait abnormalities of individuals with depression in the time domain, then we perform frequency domain analysis on this model to extract the TF-feature. In the extraction of SG-feature, we encode the positional and angular information of joints in a covariance-based descriptor to obtain the correlations between joints and utilize the metric on the symmetric positive definite (SPD) space to extract gait spatial information. Finally, two types of features are fused and fed into the prediction step for depression detection. The experimental results demonstrate that our method outperforms all other methods considered for comparison with an average accuracy of 93.75%. The main contributions of this study include the following:

- It innovatively builds a pseudo-velocity model to analyze the gait abnormalities of patients with depression during walking; the subsequently obtained TF-feature can well support depression detection.
- It provides an enhanced gait spatial geometric feature to capture complex and irregular gait patterns associated with depression, leveraging the covariance matrix and the symmetric Stein divergence on the SPD space.
- It provides a gait assessment framework based on 3D skeleton data collected by Kinect, which will motivate the development of more intelligent or simplified applications based on depth sensors for use in automatic mental health assessments.

The remainder of this paper is organized as follows. In section II, we introduce the gait data acquisition and describe the proposed gait assessment framework in detail. The experimental results are reported in section III. Finally, we conclude this paper in section IV.

II. Materials and Methods

This section mainly focuses on the construction of the Kinect-based gait assessment framework for depression detection and introduces the gait data collection. The entire framework is shown in Fig. 1.

A. Gait Data Collection

This study involved gait sequences of 95 individuals. The participants were first-year students among 2018 graduate students from various faculties of Lanzhou University, and

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

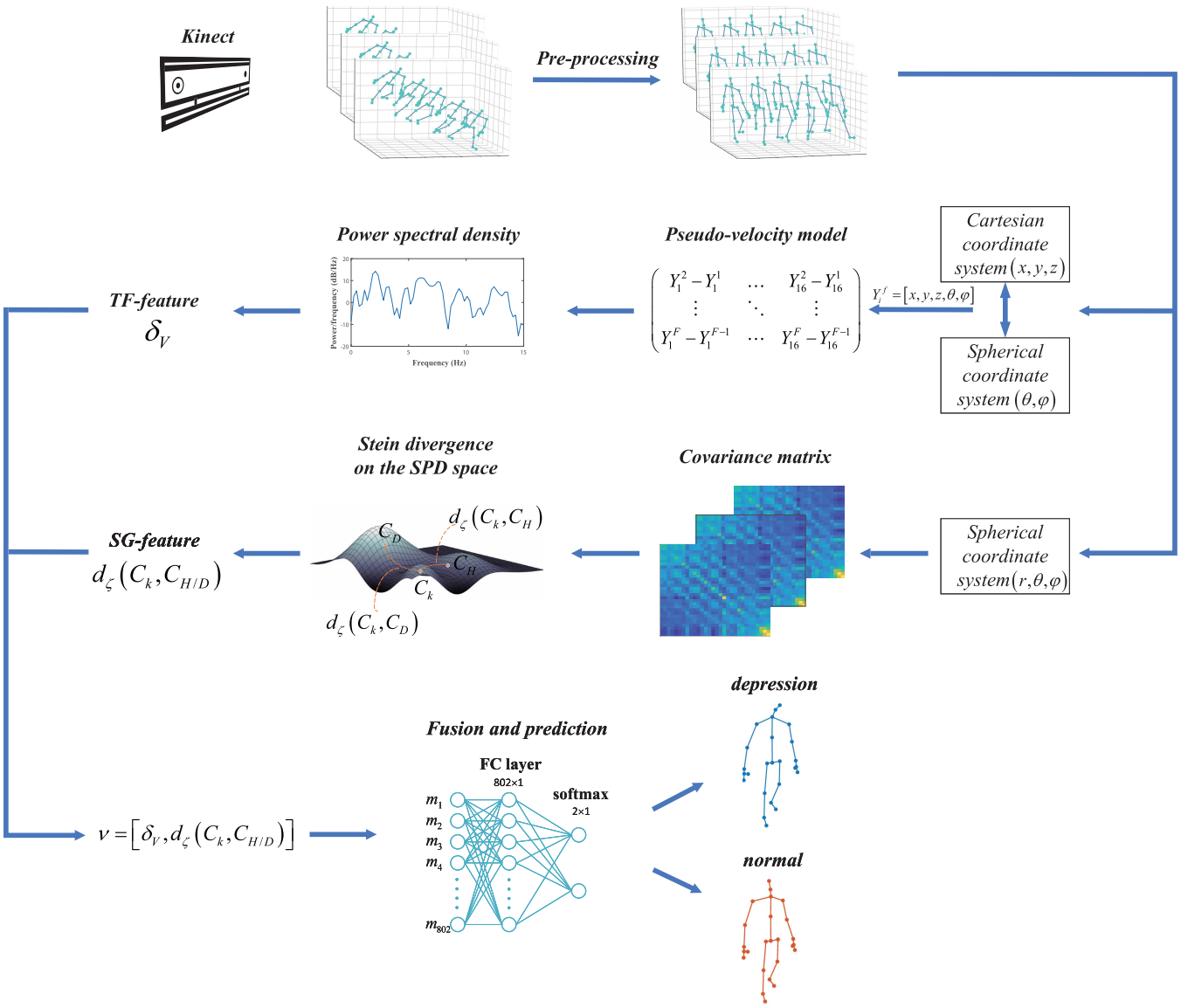


Fig. 1. The framework of our proposed method: The 3D skeleton sequences collected by Kinect are first preprocessed including resampling, angle transformation, normalization and denoising. Subsequently, the TF-feature and the SG-feature are extracted. In the extraction of the TF-feature, the pseudo-velocity model is established by combining the velocity curve in the Cartesian coordinate system with the angular velocity curve in the spherical coordinate system. Then, the power spectral density (PSD) is utilized to analyze the pseudo-velocity model for obtaining the TF-feature. In the extraction of the SG-feature, we first map the raw data into a spherical coordinate system to obtain the enhanced angular information. Then, we encode the skeleton positional and angular data by calculating the covariance matrix, after which we utilize the S-divergence to obtain the SG-feature in the SPD space. The final predictions are obtained from a softmax layer, the input of which consists of the fused features output by the fully connected (FC) layer.

their ages ranged from 22 to 28 years old. The study was approved by the Local Research Ethics Committee, and written informed consent was obtained from all participants before the experiment began. All subjects were divided into two groups based on the scores of two scales: the Patient Health Questionnaire (PHQ-9, Chinese version) [26] and Zung Self-rating Depression Scale (SDS, Chinese version) [27].

The non-depressed group (PHQ: mean = 2.04, SD = 2.72; SDS: mean = 36.85, SD = 7.21) consisted of 52 individuals (M: 28, F: 24), whose scores were in the healthy range of both the PHQ-9 and SDS. The scored-depressed group (PHQ: mean = 13.04, SD = 4.31; SDS: mean = 62.73, SD = 7.34) contained 43 (M: 23, F: 20) participants. The criteria for recruiting the scored-depressed group were that the candidates must be assessed as depression in both the PHQ-9 and SDS. In addition, the score had to reach a moderate level in at least one of them (PHQ-9 score $\geqslant$ 10 or SDS score $\geqslant$ 60). The school psychologists have confirmed the results. Basic information about the subjects is provided in Table I.

During the data collection process, all participants were required to walk in a manner that entailed two round-trip walks on a 10-meter path. Our setting for gait data collection was similar to that of [24]. Two Microsoft Kinect V2 sensors were positioned facing each other (4 meters apart) and installed on the ceiling, and their tilt angles were set to a fixed value, i.e., $-27°$ toward the path. The cameras recorded the skeleton

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

TABLE I
BASIC INFORMATION OF THE SCORED-DEPRESSED AND NON-DEPRESSED GROUPS

	Scored-depressed	Non-depressed
Cases(n)	43	52
Sex, M:F	23:20	28:24
PHQ-9 (means ± S.D)	13.04±4.31	2.04±2.72
SDS (means ± S.D)	62.73±7.34	36.85±7.21

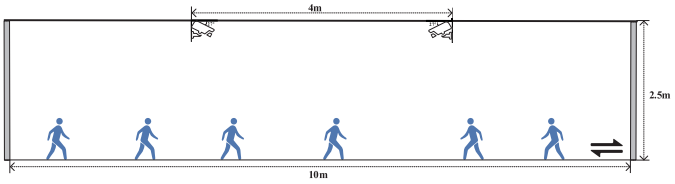


Fig. 2. The data collection scheme. Two Kinect V2 (4 meters apart) are set up in the middle of the ceiling with a −27° tilt angle toward the footpath, and the distance between the ceiling and the footpath is 2.5 meters. All participants are required to perform two round-trip walks on a 10-meter footpath.

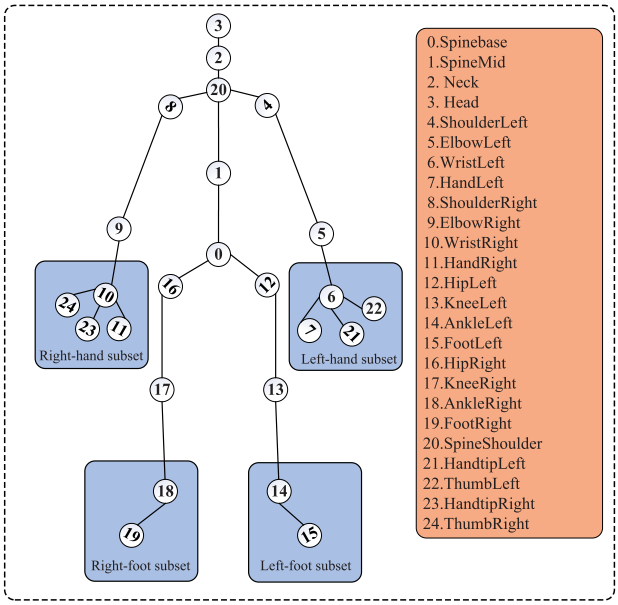


Fig. 3. The 25 markers on the human skeleton generated by Kinect V2.

joint coordinate streams at a frame rate of 30 Hz. To obtain the participants' stable and comfortable speed and posture, the gait data were collected while the subjects walked in the middle of the path. The distance between the ceiling and the ground was 2.5 meters, which served to diminish the participants' attention to the existence of the camera. To realize the simultaneous data acquisition of two Kinect cameras, we developed a gait acquisition system based on Kinect for Windows SDK 2.0 and using the C# programming language. The data collection scheme is shown in Fig. 2.

B. Data Preprocessing

1) Resampling and Angle Transformation: The skeleton streams used in the experiment are composed of 3D information from 25 joints of the entire body; the indexes of the joint points are shown in Fig. 3. To increase the measurement precision and remove unnecessary joints, we calculated the average value of the left wrist (joint 6) and the other three joints (7, 21, 22) in the left-hand subset and averaged the values of the left ankle (joint 14) and left foot (joint 15). Similar processes were applied to the right arm and leg as well, with the total number of joints decreasing from 25 to 17.

Two Kinect devices were attached to the ceiling to acquire the gait data. When participants were walking, the skeletal data from the front and back views were recorded simultaneously. Therefore, we segmented the data into back and front information based on whether the participants were facing the camera. Previous reports confirmed the effectiveness of using the skeletons from the front view over using those from the back view [24]; thus, we eliminated the segments from the back view. To ensure that each segment covered one to two gait cycles and was adapted to our subsequent feature selection, we kept the fragments that contained at least 65 frames (approximately 2.17 s). After resampling, we obtained 274 gait sequences such that the non-depressed group contained 136 segments, and the scored-depressed group contained 138 segments.

The angle of 27° between our camera located on the ceiling and the horizontal plane affected the y- and z-axis of each joint. Therefore, we performed coordinate transformation to eliminate the influence of the camera tilt. The trajectory of the i^{th} joint in the f^{th} frame after angle transformation is denoted as follows:

$$\bar{P}_i^f = \begin{bmatrix} x_i^f \\ y_i^f \\ z_i^f \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \quad f \in \mathbb{N} \tag{1}$$

where θ is equal to 27°, $i \in [1, 17]$. The results obtained before and after the angle transformation are shown in Fig. 4.

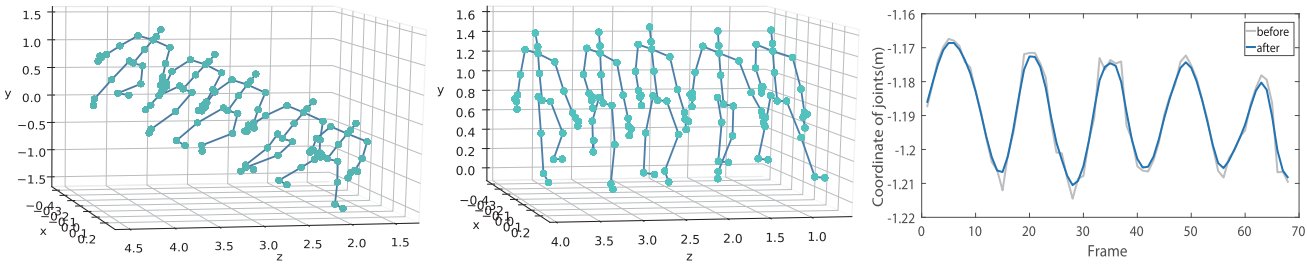
2) Normalization and Denoising: Different participants had different body sizes and positions relative to the Kinect camera. Therefore, the skeleton data required normalization to avoid errors caused by the individual differences. We used the spinebase (joint 0) as the origin of the local coordinate system and subtracted the coordinates of the spinebase from the coordinates of each joint point for each frame as follows:

$$\widetilde{P}_i^f = \begin{bmatrix} x_i^f - x_{spinebase}^f \\ y_i^f - y_{spinebase}^f \\ z_i^f - z_{spinebase}^f \end{bmatrix}, \quad f \in \mathbb{N} \tag{2}$$

where $\widetilde{P}_i^f$ is the position of the i^{th} joint in the f^{th} frame after normalization and $i \in [1, 17]$, $x_{spinebase}^f$, $y_{spinebase}^f$, $z_{spinebase}^f$ are the 3D coordinates of the spinebase in the f^{th} frame. After normalization, the 3D coordinates of the spinebase all became 0 and thus were deleted, bringing the total number of joints to 16.

In the end, we used a Gaussian filter to smooth the raw data in each dimension. It was implemented in MATLAB, with a

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c)

Fig. 4. Angle transformation and denoising results. (a) Before angle transformation. (b) After angle transformation. (c) The y-axis of the head joint in the time domain before and after Gaussian filtering.

sliding window length of 4. Fig. 4(c) provides an example of the data before and after filtering.

C. TF-Feature Extraction

In [24], the scored-depressed individuals were shown to have some gait abnormalities relative to non-depressed groups and a reduced walking speed was found to be an important and significant feature [28]. In addition to the walking speed, some studies found that the more often observed gait anomalies of individuals with depression were a slower cadence and a lethargic walking style, and even motionless and unresponsive situations [9], [29].

The slower speed, lower cadence and slumped posture reflected by individuals with depression indicate that they have motor retardation to some extent. This motor retardation may be reflected by not only the walking speed of patients with depression but also the lower speed of various parts of the body. In [30], the authors believed that the time variable exhibited by the limb velocity underlay both emotional perception and expression. Recent evidence [16] suggested that the speed curve of joints calculated by the coordinate difference can represent the nature of walking. Hence, we assume that the speed variation of each joint over time can be used to describe the motor retardation exhibited by abnormal gait in patients with depression. As illustrated in Fig. 1, we utilize the velocity curve of various body parts to indicate this speed variation over time. In particular, we regard every coordinate of each joint in the 3D skeleton data as a time series and obtain the velocity curve of each joint by calculating the difference in frames. Moreover, note that some angular features can be observed in the gait abnormalities, and the angular velocity, as a new action representation, has been proved to provide additional discriminative information [31], [32]. Therefore, to better describe the speed changes of various parts of the body, we combine the velocity in Cartesian coordinates with the angular velocity information in spherical coordinates. This angular information can be obtained by:

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos \dfrac{z}{r} \\ \varphi = \arctan \dfrac{y}{x} \end{cases} \tag{3}$$

where θ represents the upward and downward pitch of the joint, φ represents the joint around the deflection, and r is the radial distance between the origin of the coordinates and the joint. Notably, the radial distance r in (3) may have redundant information with respect to the original 3D coordinates; thus, we remove it.

We also regard every angular coordinate of each joint as a time series and perform a similar frame difference to obtain the angular velocity curve. The combination of the velocity and angular velocity curve is called the pseudo-velocity model. Let the parameter of the i^{th} joint in the f^{th} frame $Y_i^f = [x, y, z, \theta, \varphi]$, where $i \in [1, 16]$ and $f \in [1, F]$. The pseudo-velocity model is defined as:

$$\mathbf{V} = \begin{pmatrix} Y_1^2 - Y_1^1 & \cdots & Y_{16}^2 - Y_{16}^1 \\ \vdots & \ddots & \vdots \\ Y_1^F - Y_1^{F-1} & \cdots & Y_{16}^F - Y_{16}^{F-1} \end{pmatrix}, \quad Y \in \mathbb{R}^5 \tag{4}$$

Previous studies have found that in the processing of gait data, the use of time-frequency features yields better results [33]. Therefore, to analyze gait data more comprehensively, after obtaining the speed variation in various body parts over time in the time domain, we extract the frequency-domain features from the pseudo-velocity model, referred to here as the TF-feature. The PSD is one of the most important frequency-domain features, showing the strength of the energy as a function of the frequency. It has been applied to the classification of neurodegenerative diseases based on gait [34].

In this study, the spectra are calculated via the periodogram method using a 128-point FFT and periodic Hamming windows. We perform the same PSD analysis on each dimension of the pseudo-velocity model. For each of the dimensions of feature a_n, the periodogram is defined as:

$$P(f) = \frac{1}{N} \left| \sum_{n=1}^{N} a_n e^{-j2\pi fn} \right|^2 \tag{5}$$

where $n \in [1, N]$, with N being the sampling segment length (measured in frames). Given that some features are uninformative and redundant for classification, we apply the Pearson correlation, a commonly used feature selection method, to select the 10 features with the largest absolute value of the correlation coefficient on each dimension. For each segment, we obtain an $800(16 \times 5 \times 10)$-dimensional feature set δ_V.

The pseudo-velocity model considers a sequence of joints as a single time sequence, which could ignore the correlations

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

between different joints. Therefore, we also encode the positional and angular information of the joints with the covariance descriptor to obtain the correlations between joints, and we utilize the metric on the SPD space to extract the gait spatial information.

D. SG-Feature Extraction

Over the past few years, many studies have suggested that the covariance matrix has a broad range of applications based on 3D skeleton data, including person identification [35] and action recognition [36]. Recently, covariance matrices were shown to be interesting features for gait-based disease detection [8], [37]. A driving force behind this trend is that the relationship between joints reflected by the covariance matrix plays an important role in feature extraction for skeleton-based recognition tasks and is invariant to the pose and walk cycles [25], [36]. Furthermore, the resulting descriptors calculated by the covariance matrix are SPD matrices, which ensures the metrics on Riemannian manifold can be utilized to analyze the covariance matrices [13], [38]. Therefore, in this paper, we utilize covariance-based descriptors to encode the joints' positional and angular information and utilize the metric on the SPD space to extract gait spatial information.

Before encoding skeleton data through the covariance matrix, we first map the skeleton information from Cartesian coordinates into spherical coordinates. The raw information contained in the position of the 3-dimensional skeleton is limited and susceptible to the influence of individual posture; however, the angle information can avoid these problems [39]. Moreover, in [40], the authors suggested that the same data could be easily separated with a straight line in the polar coordinate system, but the situation was complicated in Cartesian coordinates. Therefore, the raw data are mapped from the rectangular coordinate system to the spherical coordinate system to obtain the extended angular information following (3). Next, we encode the skeletal joint positions and angles using the covariance matrix, which computes the correlation between the spatial information of different joints. Let $\mathbf{x}\left(\mathbf{x} \in \mathbb{R}^N\right)$ be an N-dimensional feature vector containing information about the positions and angles of 16 joints (after preprocessing, the number of joints was reduced to 16), with $N = 48(3\times16)$. The skeletal sequence mapped to the spherical coordinate system is defined as $\mathbf{X} = [x_1, \ldots, x_F] \in \mathbb{R}^{N\times F}$. F is the number of frames in the skeletal segment. Thus, the covariance matrix of the skeletal joint segment $\mathbf{X}$ is defined as:

$$\mathrm{Cov}(\mathbf{X}) = \frac{1}{F-1}\sum_{f=1}^{F}\left(x_f - \boldsymbol{\mu}\right)\left(x_f - \boldsymbol{\mu}\right)^{\mathrm{T}} \tag{6}$$

where μ is the mean of x_f, with $f \in [1, F]$.

A nonsingular covariance matrix belongs to the set of SPD matrices, which forms a connected Riemannian manifold [41]. According the Riemannian manifold theory of SPD matrices, with appropriate metrics, the geometric mean of each class can reflect the sample distribution of the class [42]. Furthermore, by taking the Bregman divergence as the metric on the SPD space, Harandi *et al.* [43] verified that the distance between two covariance matrices could be utilized by simple classifiers to obtain good classification results in action recognition based on 3D skeleton data. Therefore, in the SPD space, we assume that, according to the distance between each sample and the centers of two classes (depression and normal), the detection of depression can be performed to a certain extent. According to Algorithm 1, we exploit the training set data to obtain the class centers of the non-depressed and depressed groups, which are denoted as C_H and C_D respectively.

Algorithm 1 Method for Calculating the Center Covariance Matrices

Input: The set of covariance matrices of the same class C_n

Output: the center covariance matrix

initialization: $Center = \frac{1}{N}\sum_{n=1}^{N} C_n$,

crit= 10^8, tolerance = 10^{-8}, iteration = 0, maxiter = 50;

while *crit > tolerance and iteration < maxiter* **do**

  iteration = iteration + 1

  $C_{new} = \left[\frac{1}{N}\sum_{n=1}^{N}\left(\frac{C_n + Center}{2}\right)^{-1}\right]^{-1}$

  $crit = \|C_{new} - Center\|_F$

  $Center = C_{new}$

end

return the center of the covariance matrix $Center$;

In Algorithm 1, tolerance is the requirement for convergence stability, and maxiter is the maximum number of iterations. Then, we calculated the distance d_ζ from each sample segment to C_H and C_D, leveraging the symmetrized Bregman divergence, namely the Stein divergence [44]. The distance d_ζ is calculated by:

$$d_\zeta\left(C_k, C_{H/D}\right) = \sqrt{\log\left(\frac{\left|\left(C_k + C_{H/D}\right)/2\right|}{\sqrt{\left|C_k C_{H/D}\right|}}\right)} \tag{7}$$

where C_k is the covariance matrix of the k^{th} sample. The distances $d_\zeta\left(C_k, C_{H/D}\right)$, as the SG-feature, are fused with the TF-feature and sent to the prediction step.

E. Fusion of Features and Prediction

Given that depression consists of subtle gait deficiencies with very complex patterns, taking only a single feature is insufficient for depression detection. Therefore, we fuse the TF-feature δ_V with the SG-feature $d_\zeta\left(C_k, C_{H/D}\right)$. The fused features are defined as follows:

$$\mathcal{V} = \left[\delta_V, d_\zeta\left(C_k, C_{H/D}\right)\right] \tag{8}$$

The TF-feature and SG-feature are concatenated to obtain the 802-dimensional vector $\mathcal{V}$, which we then feed into a single fully connected (FC) layer with 802 dimensions. The FC layer can project the input of two types of features into the space of other dimensions, which may generate a new feature that is easier to separate. Following the FC layer is a softmax layer, from which the final predicted probability value is obtained. The softmax function projects the feature of each sample segment onto a scale of the same size, and

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

the predicted category has a greater probability value. The structure of the fusion and prediction is shown in Fig. 1.

III. Results and Discussion

To verify the validity of our method, we design two experiments. We first analyze the statistical results of the pseudo-velocity model and the rationality of feature fusion using our methods. Then, we conduct a comparative analysis of the current representative methods using our data, which also used the covariance matrix and the time and frequency characteristics to analyze Kinect-captured gait data.

We conduct experiments on the depression gait dataset of postgraduate students described in section II-A. To evaluate the performance of the classifiers, we adopt a 'leave-persons-out' protocol, keeping all segments of a person as the test set and the other segments as the training set. Since some important features with small values or negative values may be ignored when training the model, we use the min-max normalization method to normalize the features into $[-1, 1]$ before the classification step [45]. In the fusion and prediction step, we use a single FC layer with layer size of 802 hidden nodes. The activation function is ReLU and the loss is estimated as the cross-entropy loss. The optimizer is Adam with a learning rate of 1e-5. In addition, the early stopping trick is applied to avoid overfitting. The accuracy, sensitivity, specificity, F1-score, and the area under the curve (AUC) of the receiver operating characteristic (ROC) are used to evaluate the experimental performance.

A. Experimental Results of TF-Feature and SG-Feature

1) Data Analysis for the Pseudo-Velocity Model: In section II-C, we build a pseudo-velocity model to analyze the gait abnormalities. To intuitively illustrate that the pseudo-velocity model can represent the slower cadence and motor retardation of patients with depression, we analyze the means and the standard deviations of the pseudo-velocity model on four representative joints in our data, as shown in Table II. To show the difference in the pseudo-velocity model between the scored-depression and non-depressed groups more clearly, we list the statistical results of the pseudo-velocity model on the 4 joints with the greater amounts of activity in the table, including the right elbow, right wrist, right hip, and the right knee. Without loss of generality, we average the absolute values of the velocity curves and angular velocity curves of the relevant joints in all segments of the two classes.

As shown in Table II, the averages of the variations in the velocity and angular velocity of the scored-depressed individuals are generally slower than those of the individuals without depression during gait. In addition, the scored-depression group clearly has a smaller standard deviation, which shows that the range of changes in the velocity and angular velocity of their related joints is not sufficiently dramatic. The results suggest that the pseudo-velocity model can describe the motor retardation and slower cadence of scored-depressed individuals to a certain extent and has potential value in the characterization of depression.

TABLE II
THE STATISTICAL RESULTS OF THE PSEUDO-VELOCITY MODEL IN THE SCORED-DEPRESSION (SD) AND NON-DEPRESSED (ND) GROUPS. WE AVERAGE THE VELOCITY CURVES AND ANGULAR VELOCITY CURVES OF THE RELEVANT JOINTS FROM ALL SEGMENTS IN EACH CLASS. WE LIST THE FOUR JOINTS ON THE LIMBS WITH GREATER ACTIVITY, I.E., THE RIGHT ELBOW (R.ELBOW), RIGHT WRIST (R.WRIST), RIGHT HIP (R.HIP), AND RIGHT KNEE (R,KNEE)

		x	y	z	θ	φ
R.Elbow	SD	0.51±0.37	0.98±0.65	1.04±0.76	0.26±0.19	0.13±0.10
	ND	0.53±0.48	1.19±0.85	1.36±1.05	0.31±0.24	0.17±0.14
R.Wrist	SD	0.53±0.34	1.47±0.79	1.80±1.24	0.39±0.28	0.31±0.19
	ND	1.05±0.70	1.85±1.30	2.55±1.91	0.58±0.42	0.38±0.26
R.Hip	SD	0.14±0.13	0.25±0.19	0.18±0.15	0.15±0.12	0.19±0.14
	ND	0.15±0.13	0.32±0.23	0.21±0.19	0.19±0.15	0.24±0.18
R,Knee	SD	0.89±0.66	1.62±1.20	2.64±1.70	1.43±0.77	0.28±0.20
	ND	1.10±0.87	1.91±1.35	3.32±1.51	1.78±1.29	0.31±0.31

TABLE III
COMPARISON OF THE FEATURE CLASSIFICATION RESULTS (IN PERCENTAGE) BEFORE AND AFTER FUSION IN TERMS OF FOUR EVALUATION CRITERIA (ACCURACY, SENSITIVITY, SPECIFICITY, AND F1-SCORE). THE BEST RESULTS IN THIS TABLE ARE LABELED IN BOLD

	Accuracy	Sensitivity	Specificity	F1
Velocity	88.84±7.13	90.08±5.99	87.38±6.64	89.71±12.46
Angle Velocity	83.93±7.33	83.47±6.49	84.47±5.91	84.87±14.41
TF(Pseudo-Velocity Model)	91.96±5.33	91.74±4.28	92.23±4.45	92.50±9.33
SG	70.98±11.85	72.73±11.64	68.93±11.12	73.03±18.74
TF+SG	**93.75**±2.98	**93.39**±2.20	**94.17**±2.25	**94.17**±6.59

2) Feature Fusion Results: The TF-feature is derived from the pseudo-velocity model, which is composed of the velocity curve in the rectangular space with the angular velocity curve in the spherical coordinates. To prove experimentally that the TF-feature extracted from pseudo-velocity model is more effective than that using only the velocity curve, we conduct three experiments, extracting the TF-feature from three different velocity models, namely the velocity curve, the angular velocity curve, and the pseudo-velocity model combining the first two. In addition, to confirm the effectiveness of the fused features obtained by our method, we investigate the classification performance when using only the TF-feature or SG-feature. The experimental results are shown in Table III. When we use only the TF-feature extracted from the three different velocity models for classification, instead of fusing it with the SG-feature, the TF-feature is directly fed into the FC layer. When using only the SG-feature, the method is similar to that above.

We first compare the classification results using three different velocity models. When the TF-feature is extracted from pseudo-velocity model, the classification accuracy can reach 91.96%, significantly outperforming the cases in which only the velocity curve or only the angular velocity is used, and the sensitivity, specificity and F1-score improve significantly.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

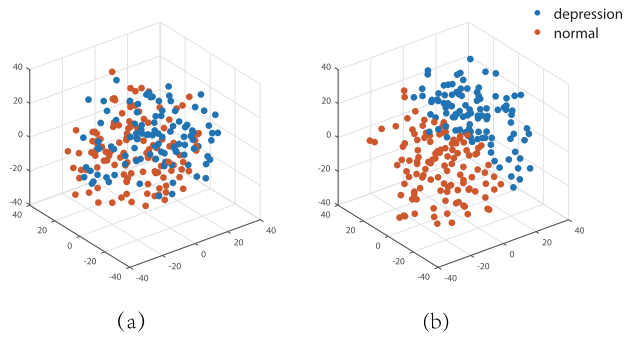


Fig. 5. A t-SNE dimension-reduction visualization map. Each point represents a segment of the gait skeleton data. The two colors represent different classes, with data points belonging to the same class having the same color. (a) 3D perspective of the raw data before feature extraction. (b) 3D perspective of the fused features.

These results prove that combining the velocity curve with the angular velocity curve, i.e., constructing the pseudo-velocity model, is necessary and correct. Then, we compare the results before and after the fusion of the TF-feature and SG-feature. As shown in Table III, the classification average accuracy achieved by using the SG-feature alone is 70.98%. By contrast, the accuracy using the TF-feature alone can reach 91.96%, but compared with the fused feature, its performance is lower by approximately 1.79%. Furthermore, the sensitivity, specificity and F1-score of the fused feature are also improved by 1.65%, 1.94% and 1.67% respectively. The smaller standard deviations also indicate that the fused features can effectively alleviate the impact of individual differences. This finding indicates that the TF-feature extracted from the pseudo-velocity model is a more discriminative feature for classification. Furthermore, the incorporation of the SG-feature can yield better classification performance.

To more intuitively illustrate whether the fused features are separable, we visualize the skeleton data before feature extraction and the fused features through the t-distributed stochastic neighbor embedding (t-SNE). The t-SNE algorithm can reduce the high-dimensional data to a relatively low-dimensional subspace. In this paper, the data are visualized in a 3D space, as shown in Fig. 5. The visualization map clearly shows that the two categories of data without feature extraction are mixed haphazardly, while the features after fusion are separable in the 3D space. This outcome further illustrates that the fused features have an enhanced expression ability.

B. Comparison With Different Methods

To investigate the performance advantages obtained by the fused features, we also compared our method with several similar gait research methods. Because the methods of gait-based depression detection have not yet been well established, the methods that we compared included the use of time-frequency characteristics and covariance matrices for gait-based emotion recognition and other types of disease analysis. Leveraging the covariance matrices to encode joints' positional and speed data, a method for classification of gait disorders arising from Parkinson's and hemiplegia was presented in [8]. In [16], the authors proposed a method using

TABLE IV
COMPARISON OF THE CLASSIFICATION RESULTS (IN PERCENTAGE) BETWEEN OUR METHOD AND OTHER METHODS IN TERMS OF VARIOUS EVALUATION CRITERIA (ACCURACY, SENSITIVITY, SPECIFICITY, AND F1-SCORE) AND INDICATIONS OF SIGNIFICANT DIFFERENCES. THE BEST RESULTS IN THIS TABLE ARE LABELED IN BOLD

	Accuracy	Sensitivity	Specificity	F1
Li et al. [8]	68.30±12.15**	72.73±11.49**	63.11±11.96**	71.26±18.31**
Li et al. [16]	71.43±14.23**	74.38±14.72**	67.96±14.51**	73.77±19.85**
Yuan et al. [23]	85.71±7.52**	85.95±6.79*	85.44±6.50*	86.67±13.78**
Fang et al. [24]	91.58±7.26*	86.05±6.90*	**96.15**±8.67	90.24±7.15*
Our method	**93.75**±2.98	**93.39**±2.20	94.17±2.25	**94.17**±6.59

(*: $p \leq 0.05$, **: $p \leq 0.01$)

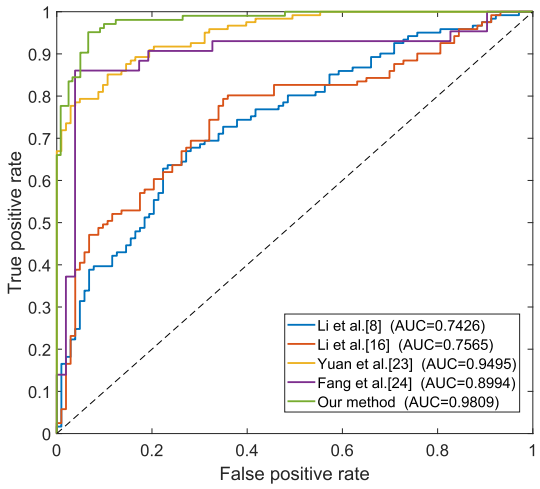


Fig. 6. The ROC curves between the proposed method and other methods.

the discrete Fourier transform and statistical methods to extract various time-frequency features to recognize emotion from gait information. The authors of [23] proposed a method to detect depression based on the Kinect-captured gait data, and the researchers utilized the Hilbert-Huang transform to extract gait frequency features. In [24], the authors applied abnormal gait features in the time domain to identify depression. Table IV reports the averages, standard deviations, and indications of significant differences among various metrics between the proposed method and other methods.

As shown in Table IV, it can be easily observed that the mean accuracy of the proposed approach outperforms other methods, achieving state-of-the-art results. Furthermore, it should be noted that the standard deviations of our method are less than that with other methods, which is indicative of the robustness of the proposed method, emphasizing that the results are not affected by the test data to a great extent. We find that there are significant differences between the experimental results with our method and those with other methods using the paired sample t-test. This reveals that the proposed method corresponds to a significant improvement over other methods. The AUC is known to well validate

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

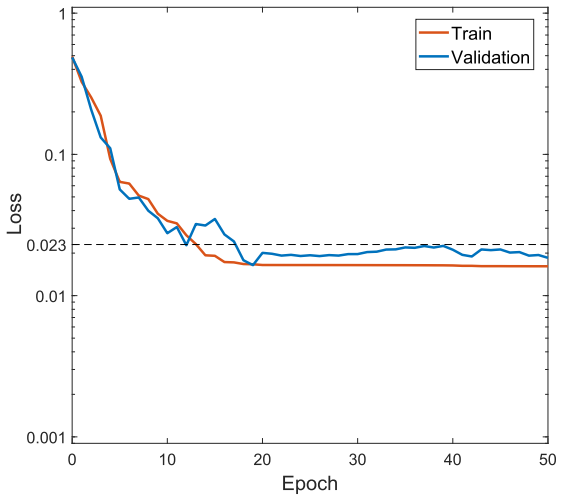


Fig. 7. The training and validation losses curve of the proposed method.

the performance of binary classification models [46], so we utilize it to test the discriminability of the models. Fig. 6 illustrates the ROC curve of our approach and other methods. The AUC of the proposed method is the highest reported, which characterizes the robustness and stability of our method for the detection of depression. Fig. 7 presents the training and validation losses curve of our method. In both the training and validation sets, total loss converge to 0.023 after 18 epochs. This result further proves that our method has convergence. In general, by fusing both TF-feature and SG-feature, our method can alleviate the impact of individual differences and enhance the classification robustness and performance.

IV. Conclusion

Gait analysis offers a new low-cost and contactless method for depression detection. Due to the development of inexpensive and portable depth sensors such as Microsoft Kinect, the advantages of gait analysis in the detection of depression are further amplified. Therefore, in this paper, a novel framework is proposed for implementing non-intrusive, real-time and automatic depression detection from the gait data collected by Microsoft Kinect. In contrast to previous methods that focused only on specific abnormal gait characteristics, we innovatively propose a pseudo-velocity model to describe the slower cadence and motor retardation exhibited by gait abnormalities in individuals with depression; the subsequently extracted TF-features can well support depression detection. In addition, to capture complex and irregular gait spatial information, we encode joint positional and angular information in a covariance-based descriptor, utilizing the S-divergence on the SPD space to obtain the SG-feature. The new features obtained after the fusion of these two types of features are effective in characterizing and indicating depression. The proposed gait assessment framework for depression detection using Kinect could motivate more intelligent, convenient and objective applications based on depth sensors in the field of automatic mental health.

References

[1] World Health Organization. *Depression Fact Sheets, WHO*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[2] A. Halfin, "Depression: The benefits of early and appropriate treatment," *Amer. J. Managed Care*, vol. 13, no. 4, pp. 92–97, Nov. 2007.

[3] H. Cai *et al.*, "A pervasive approach to EEG-based depression detection," *Complexity*, vol. 2018, Feb. 2018, Art. no. 5238028.

[4] H. Jiang *et al.*, "Investigation of different speech types and emotions for detecting depression using different classifiers," *Speech Commun.*, vol. 90, pp. 39–46, Jun. 2017.

[5] M. Trotzek, S. Koitka, and C. M. Friedrich, "Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 3, pp. 588–601, Mar. 2020.

[6] Z. Yao *et al.*, "Morphological changes in subregions of hippocampus and amygdala in major depressive disorder patients," *Brain Imag. Behav.*, vol. 14, no. 3, pp. 653–667, Jun. 2020.

[7] P. Sawangjai, S. Hompoonsup, P. Leelaarporn, S. Kongwudhikunakorn, and T. Wilaiprasitporn, "Consumer grade EEG measuring sensors as research tools: A review," *IEEE Sensors J.*, vol. 20, no. 8, pp. 3996–4024, Apr. 2020.

[8] Q. Li *et al.*, "Classification of gait anomalies from kinect," *Vis. Comput.*, vol. 34, no. 2, pp. 229–241, Feb. 2018.

[9] M. R. Lemke, T. Wendorff, B. Mieth, K. Buhl, and M. Linnemann, "Spatiotemporal gait patterns during over ground locomotion in major depression compared with healthy controls," *J. Psychiatric Res.*, vol. 34, nos. 4–5, pp. 277–283, Jul. 2000.

[10] J. Michalak, N. F. Troje, J. Fischer, P. Vollmar, T. Heidenreich, and D. Schulte, "Embodiment of sadness and depression—Gait patterns associated with dysphoric mood," *Psychosomatic Med.*, vol. 71, no. 5, pp. 580–587, Jun. 2009.

[11] J. Z. Canales, T. A. Cordás, J. T. Fiquer, A. F. Cavalcante, and R. A. Moreno, "Posture and body image in individuals with major depressive disorder: A controlled study," *Revista Brasileira de Psiquiatria*, vol. 32, no. 4, pp. 375–380, Dec. 2010.

[12] N. M. DiFilippo and M. K. Jouaneh, "Characterization of different microsoft kinect sensor models," *IEEE Sensors J.*, vol. 15, no. 8, pp. 4554–4564, Aug. 2015.

[13] C. Tang, W. Li, P. Wang, and L. Wang, "Online human action recognition based on incremental learning of weighted covariance descriptors," *Inf. Sci.*, vol. 467, pp. 219–237, Oct. 2018.

[14] J. Shotton *et al.*, "Real-time human pose recognition in parts from single depth images," in *Proc. CVPR*, Jun. 2011, pp. 1297–1304.

[15] B. Ghojogh, H. Mohammadzade, and M. Mokari, "Fisherposes for human action recognition using kinect sensor data," *IEEE Sensors J.*, vol. 18, no. 4, pp. 1612–1627, Feb. 2018.

[16] B. Li, C. Zhu, S. Li, and T. Zhu, "Identifying emotions from non-contact gaits information based on microsoft kinects," *IEEE Trans. Affect. Comput.*, vol. 9, no. 4, pp. 585–591, Oct. 2018.

[17] S. Bei, Z. Zhen, Z. Xing, L. Taocheng, and L. Qin, "Movement disorder detection via adaptively fused gait analysis based on kinect sensors," *IEEE Sensors J.*, vol. 18, no. 17, pp. 7305–7314, Sep. 2018.

[18] A. González, P. Fraisse, and M. Hayashibe, "Adaptive interface for personalized center of mass self-identification in home rehabilitation," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2814–2823, May 2015.

[19] M. Parajuli, D. Tran, W. Ma, and D. Sharma, "Senior health monitoring using kinect," in *Proc. 4th Int. Conf. Commun. Electron. (ICCE)*, Aug. 2012, pp. 309–312.

[20] F. Gholami, D. A. Trojan, J. Kovecses, W. M. Haddad, and B. Gholami, "A microsoft kinect-based point-of-care gait assessment framework for multiple sclerosis patients," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 5, pp. 1376–1385, Sep. 2017.

[21] P. Ren *et al.*, "Multivariate analysis of joint motion data by kinect: Application to Parkinson's disease," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 181–190, Jan. 2020.

[22] N. Zhao *et al.*, "See your mental state from your walk: Recognizing anxiety and depression through kinect-recorded gait data," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0216591.

[23] Y. Yuan, B. Li, N. Wang, Q. Ye, Y. Liu, and T. Zhu, "Depression identification from gait spectrum features based on Hilbert-Huang transform," in *Proc. 4th Int. Conf. Hum. Centered Comput. (HCC)*, Dec. 2018, pp. 503–515.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

[24] J. Fang *et al.*, "Depression prevalence in postgraduate students and its association with gait abnormality," *IEEE Access*, vol. 7, pp. 174425–174437, 2019.

[25] M. S. N. Kumar and R. V. Babu, "Human gait recognition using depth camera: A covariance based approach," in *Proc. 8th Indian Conf. Comput. Vis., Graph. Image Process.*, Dec. 2012, p. 20.

[26] 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.

[27] W. W. Zung, "A self-rating depression scale," *Arch. Gen. Psychiatry*, vol. 12, no. 12, pp. 63–70, Jan. 1965.

[28] R. Briggs *et al.*, "Do differences in spatiotemporal gait parameters predict the risk of developing depression in later life?" *J. Amer. Geriatrics Soc.*, vol. 67, no. 5, pp. 1050–1056, May 2019.

[29] C. Wilkes, R. Kydd, M. Sagar, and E. Broadbent, "Upright posture improves affect and fatigue in people with depressive symptoms," *J. Behav. Therapy Exp. Psychiatry*, vol. 54, pp. 143–149, Mar. 2017.

[30] A. Barliya, L. Omlor, M. A. Giese, A. Berthoz, and T. Flash, "Expression of emotion in the kinematics of locomotion," *Exp. Brain Res.*, vol. 225, no. 2, pp. 159–176, Mar. 2013.

[31] G. T. Papadopoulos, A. Axenopoulos, and P. Daras, "Real-time skeleton-tracking-based human action recognition using kinect data," in *Proc. Int. Conf. Multimedia Modeling*, Jan. 2014, pp. 473–483.

[32] M. Karg, K. Kuhnlenz, and M. Buss, "Recognition of affect based on gait patterns," *IEEE Trans. Syst., Man, Cybern., B (Cybern.)*, vol. 40, no. 4, pp. 1050–1061, Aug. 2010.

[33] E. Sejdic, K. A. Lowry, J. Bellanca, M. S. Redfern, and J. S. Brach, "A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 3, pp. 603–612, May 2014.

[34] K. Devi Das, A. J. Saji, and C. S. Kumar, "Frequency analysis of gait signals for detection of neurodegenerative diseases," in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Apr. 2017, pp. 1–6.

[35] A. Wu, W.-S. Zheng, and J.-H. Lai, "Robust depth-based person re-identification," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2588–2603, Jun. 2017.

[36] M. E. Hussein, M. Torki, M. A. Gowayyed, and M. El-saban, "Human action recognition using a temporal hierarchy of covariance descriptors on 3D joint locations," in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, Aug. 2013, pp. 2466–2472.

[37] E. Dolatabadi, A. Mansfield, K. K. Patterson, B. Taati, and A. Mihailidis, "Mixture-model clustering of pathological gait patterns," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 5, pp. 1297–1305, Sep. 2017.

[38] L. Wang, J. Zhang, L. Zhou, C. Tang, and W. Li, "Beyond covariance: Feature representation with nonlinear kernel matrices," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4570–4578.

[39] T.-N. Nguyen, H.-H. Huynh, and J. Meunier, "Skeleton-based abnormal gait detection," *Sensors*, vol. 16, no. 11, p. 1792, Oct. 2016.

[40] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[41] O. Tuzel, F. Porikli, and P. Meer, "Human detection via classification on Riemannian manifolds," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1297–1304.

[42] M. Congedo, A. Barachant, and R. Bhatia, "Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review," *Brain-Comput. Interfaces*, vol. 4, no. 3, pp. 155–174, Mar. 2017.

[43] M. Harandi, M. Salzmann, and F. Porikli, "Bregman divergences for infinite dimensional covariance matrices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1003–1010.

[44] S. Sra. (2011). *Positive Definite Matrices and the Symmetric Stein Divergence*. [Online]. Available: http://people.kyb.tuebingen.mpg.de/suvrit/

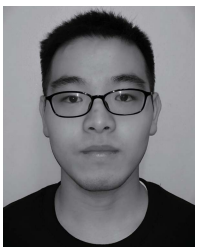
[45] A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics* (International Series on Biometrics). Springer, 2006.

[46] A. Ditthapron, N. Banluesombatkul, S. Ketrat, E. Chuangsuwanich, and T. Wilaiprasitporn, "Universal joint feature extraction for P300 EEG classification using multi-task autoencoder," *IEEE Access*, vol. 7, pp. 68415–68428, 2019.

Tao Wang (Graduate Student Member, IEEE) received the bachelor's degree from Zhengzhou University, China, in 2018. He is currently pursuing the master's degree with the Universal Perception and Intelligent Systems Laboratory, Lanzhou University, China. He is also a Guest Student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research focuses on bioelectrical signal processing and automated mental disorder detection.

Cancheng Li was born in Guangdong, China. He received the bachelor's degree in 2017. He is currently pursuing the master's degree with the Universal Perception and Intelligent Systems Laboratory, Lanzhou University, China. He is also a Guest Student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests are bioelectrical signal processing and pattern recognition. He has been a Reviewer of IEEE ACCESS Journal since 2018.

Chunyun Wu (Graduate Student Member, IEEE) received the bachelor's degree in electronic and information engineering from Fuzhou University, China, in 2017. She is currently pursuing the master's degree with the Universal Perception and Intelligent Systems Laboratory, Lanzhou University, China. She has been an Exchange Student with National Taiwan University since 2015. Her main research interests include near infrared imaging, neural networks, and bioelectrical signal processing.



Chengjian Zhao (Graduate Student Member, IEEE) received the B.E. degree from Chongqing University, Chongqing, China, in 2016. He is currently pursuing the M.E. degree in communication and information system with the School of Information Science and Engineering, Lanzhou University, Lanzhou, China. His main research interests include pattern recognition, neural networks, intelligent information processing, artificial intelligence, and bioelectrical signal processing.

Jieqiong Sun received the bachelor's degree in electronic science and technology from the Beijing University of Posts and Telecommunications, China. She is currently pursuing the master's degree in communication and information system with the School of Information Science and Engineering, Lanzhou University, China. Her research interests focus on bioelectrical signal processing and pattern recognition.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.

Hong Peng received the Ph.D. degree from Lanzhou University, Lanzhou, China. From 2010 to 2011, he was as a Visiting Scholar with the Institute of Computer System, ETH Zurich, Switzerland. He is currently an Associate Professor with the School of Information Science and Engineering, Lanzhou University. He is also in charge of three projects from the National Natural Science Foundation of China, the Central College Foundation Project of Lanzhou University, and the Youth Cross-Project of Lanzhou University. He has authored or coauthored more than 30 papers in peer-reviewed journals, conferences, and book chapters. He is involved in the work of biosensors, biological signal processing, and emotional characteristics analysis. His research areas include bioinformation processing and ubiquitous affective computing.

Xiping Hu (Member, IEEE) received the Ph.D. degree from the University of British Columbia, Vancouver, BC, Canada. He is currently a Professor with Lanzhou University, China. He was the Co-Founder and the CTO of Bravolol Ltd., Hong Kong, a leading language learning mobile application company with more than 100 million users, and listed as the top two language education platform globally. He has more than 90 papers published and presented in prestigious conferences and journals, such as the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING (TETC)/IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS (TII)/IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (TVT)/IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (TITS)/IEEE INTERNET OF THINGS (IoT) JOURNAL/IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC), ACM TOMM, the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (COMST), the *IEEE Communications Magazine*, the IEEE NETWORK, HICSS, ACM MobiCom, AAAI, and WWW. He has been serving as the lead guest editors of the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and WCMC, and an Associate Editor of IEEE ACCESS. His research areas consist of mobile cyber-physical systems, crowdsensing, and affective computing.

Bin Hu (Senior Member, IEEE) received the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Science, Beijing. From 2009 to 2019, he served as the Dean of the School of Information Science and Engineering, Lanzhou University, Lanzhou, China. His current research interests include pervasive computing, cognitive computing, and mental health care. He served as an Editor of *IET Communications*, *Cluster Computing*, *Wireless Communications and Mobile Computing*, the *Journal of Internet Technology*, *Security and Communication Networks* (Wiley), and *Brain Informatics*, and an Associate Editor of some peer-reviewed journals in computer science.

Authorized licensed use limited to: TIANJIN UNIVERSITY. Downloaded on August 19,2023 at 11:17:23 UTC from IEEE Xplore. Restrictions apply.