Emotion recognition based on weighted kernel support vector machine using wearable inertial sensors

Yan Zhao
Linyi University
Ming Guo (guoming0537@126.com)
Linyi University
Xiangyong Chen
Linyi University
Feng Zhao
Linyi University
Jianlong Qiu
Linyi University

Research Article

Keywords: Emotion recognition, Weighted kernel support vector machine, Seneor system, Fuzzy comprehensive evaluation

Posted Date: May 31st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1629621/v1

RESEARCH

Emotion recognition based on weighted kernel support vector machine using wearable inertial sensors

Yan Zhao, Ming Guo*, Xiangyong Chen, Feng Zhao and Jianlong Qiu

*Correspondence: guoming0537@126.com
School of Automation and Electrical Engineering, Linyi University, Linyi, China
Full list of author information is available at the end of the article

Abstract

Emotion recognition has various applications closely linked with people's daily life, such as human-computer interaction and psychotherapy. So this paper focuses on the research of emotion recognition of body movement and proposes a novel emotion recognition approach based on weighted kernel support vector machine (SVM) using wearable inertial sensors. Specifically, the mapping relationship from emotions to body movements is established by fuzzy comprehensive evaluation. The subjects wear inertial sensors on their arms and wrists to collect data in six emotions including sleepy, bored, excited, tense, anger and distressed. In recognition phase, the weighted kernel SVM model is constructed, in which the fuzzy function is auxiliary to improve the weight calculation method of kernel functions in multiple kernel SVM. It also explores the effect of different combinations of inertial sensors on the recognition. The results show that compared with other methods, the proposed method achieves 98.4% accuracy for six emotions, which is effective and applicable.

Keywords: Emotion recognition; Weighted kernel support vector machine; Seneor system; Fuzzy comprehensive evaluation

1 Introduction

Emotion is a psychological response to external stimuli or changes. When emotion comes into being, there are subjective and physiological emotional response [1]. Emotional expression mainly in facial expressions, speech and body movements, which are also the source of emotional information. Nowadays, artificial intelligence and human-computer interaction are showing a trend of rapid growth. Emotion recognition played an important role in improving the intelligence of human-computer interaction. Only when real artificial emotions are realized, artificial intelligence can have a broader space for development and form a stronger driving force for social productivity. In 1966, Reeves and Nass of Stanford University pointed out in their study that the problem needs to be solved in human-computer interaction is emotional intelligence, which is consistent with people's communication [2].

The current researches on emotion recognition are mainly realized speech signal and facial expression. Body movement, as an important part of body language, also have rich emotions. Body movement recognition is widely used in Biomedical monitoring [3], clinical evaluation, sports competition [4] and other fields. Research results show that body movement is also an important way of expressing emotions

[5]. People perceive the emotional states of other individuals through body movements in daily life, and different emotional states show different body movements. In reality, human-computer interaction scenarios are complex, emotion recognition methods based on body movement can play an important role when facial expression and speech are invalid.

Emotion recognition is an important link in emotion computing. We obtain features that can express emotion to the greatest extent from various human body signals, and find the mapping relationship between external representation and internal state. Through machine learning algorithms to identify the current emotion.

Most studies mainly use visual methods [6, 7], that can provide rich information and intuitive performance, but with some congenital defects. Image or video processing technology requires a large number of memory resources on high-specification computers, which results in high complexity. In addition, factors such as illumination and occlusion in the monitoring scene can affect the recognition accuracy.

High-performance inertial sensor systems have made a great breakthrough and have been widely used in health monitoring [8], communication positioning. It has built-in inertial measurement units (IMU) such as accelerometers, gyroscopes, and magnetometers and is worn on one or several parts of the human body for data collection. The use of inertial sensors perfectly compensates for the shortcomings of the visual method. These systems are cheaper with respect to their vision-based counterparts, can easy to collect data for a long time without being affected by the environment.

This paper mainly studies emotion recognition of body movement using wearable inertial sensors. We establish one-to-one correspondence between emotions and body movements and collected data in six emotions, including sleepy, bored, excited, tense, anger and distressed. Then, the emotional data is divided by sliding windows. For each window, we extract features in time domain and frequency domain, all of which make up the input. The weighted kernel SVM is constructed as a classifier to predict emotion categories. The innovation points of this paper are:

1. This paper proposes a new emotion recognition approach based on body movements, which uses wearable inertial sensors to acquire data and predict emotional states through weighted kernel SVM. As there are currently few related studies, which also provides a reference for follow-up research.

2.This paper presents a weight calculation method using fuzzy function in multiple kernel SVM. Based on this, the weighted kernel SVM optimization algorithm is proposed to increase the accuracy of emotion recognition.

3.The contribution of different inertial sensor combinations to the overall recognition accuracy is explored, aiming to find out the minimization of the sensor system under the premise of acceptable accuracy.

4.The results show that the proposed menthod outperform the existing methods in emotion recognition accuracy, and the recognition accuracy in six emotions is 98.4%.

2 Related Work

Facial expression is one of the direct ways to express emotions, and it is also one of the earliest researched topics in emotion recognition. Massive scholars have conducted in-depth research on facial emotion recognition, and have achieved mature

results. Alia K.Hassan et al. [9] proposed a facial emotion recognition method based on graph mining, which represented the face region as a node and edge graph and the results show that the accuracy is improved by about 2%. P.Kaviya et al. [10] used deep learning to propose a facial emotion recognition system, that used Haar filters to detect and extract facial features and recognize expressions through CNN. Austin Nicolai et al. [11] proposed an emotion recognition method based on fuzzy logic. They first used image processing technology to obtain facial regions, and then extracted feature recognition points, and used them to blur and determine the intensity of different facial actions, thereby judging emotions. Wenhui Shi et al. [12] embed wavelet neural network into a fuzzy neural network (FNN) to obtain fuzzy wavelet neural network, through the use of FNN to recognize facial emotions and used Levenberg-Marqudt algorithm to reduce training time.

Speech is an important way to express emotions. Most of the researches use acoustic features for speech emotion recognition and analyze the correlation between emotions. Jianhua Tao et al. [13] used semi-supervised learning to construct a ladder network for speech emotion recognition with a small amount of labeled data and achieved good recognition efficiency. Xiaoqing Jiang et al. [14] first eliminated the noise by reconstructing the speech signal in a noisy environment, then used the multi-core support vector machine model solved by SDP to recognize the speech signal. Jianfeng Zhao et al. [15] designed a merged convolutional neural network (CNN) consisting of a one-dimensional (1D) CNN branch and a 2D CNN branch, and used Bayesian optimization to select parameters to learn deep features from different data to recognize speech emotions. In the speech emotional feature extraction, Shaoling Jing et al. [16] proposed a prominence-related feature, and combined it with traditional acoustic features to maximize the retention of emotional information.

Wearable inertial sensors are widely used to record body movement data, such as pedestrian estimation [17], activity recognition [18], and gait evaluation [19, 20]. Recent attempts to recognize emotions from body movement include fixing wearable inertial devices to one or more positions on the subject to record emotional data [21, 22, 23]. M.A. Hashmi et al. [21] used human gait inertial signals recorded by a smartphone worn on the chest to identify six basic emotions, and used SVM and random forests as prediction models. Gravina et al. [22] use sensor-level and feature-level fusion techniques to identify and monitor activities within seats to reflect psychological and emotional states. Zhang et al. [23] tracked and recorded movements using smart bracelets worn on the tester's right wrist and ankle, which accurately identified human moods.

3 Materials and Methods

The general process of the proposed emotional recognition methods is shown in Figure 1, including the establishment of emotional dataset, feature extraction and selection, emotional model training and classification.

3.1 Emotion Dataset

To realize emotion recognition, the first step is to define emotion. In 1980, Russell proposed the emotional circle model, which is widely approved and used in various

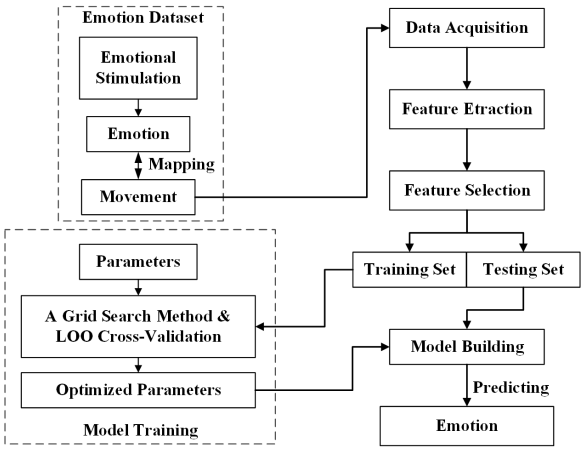

Figure 1 The process of emotion recognition based body movement using inertial sensors

studies of emotion recognition [24], as shown in Figure 2. Emotions are divided into pleasure and arousal dimensions, and some areas in the model space are interpreted as 28 discrete emotions. Then, it also needs the support of public datasets. Refer to the research of Klaus R. Scherer and HeinerEllgring [25], who listed multiple types of limb movements and 14 common emotions, and gave the correlation between them through statistical experiments.

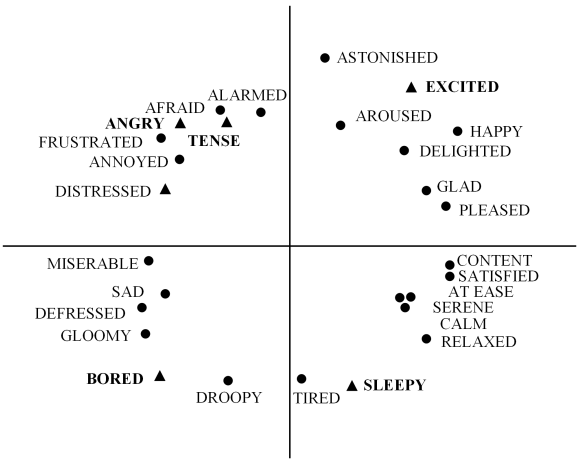

Figure 2 The circumplex model of emotion

In general, the same body movement expresses different emotions in different situations, so it is difficult to establish a corresponding relationship between emotions and body movements. Up to now, there are few related literatures and datasets available for reference [26]. On the one hand, datasets collected in specific laboratory environments are obtained by professional actors through performance. The quality of the datasets has a direct relationship with the performance of the actors. On the other hand, for non-performative datasets, different people have different emotional judgments of the same body movement, which is strong subjective.

This paper establishes an emotion dataset composed of body movement labels and emotion labels, based on the 14 basic emotions mentioned by [25], some of which are not involved. Initially summarized sleepy, bored, excited, tense, anger and distressed as emotional labels, which are marked as triangle in Figure 2. Body movement labels include shake the fist, applaud, cover the face with both hands, rub hands, wave hands, and stretch.

It is not enough to only refer to the datasets provided in the relevant literature, but also to determine the mapping relationship between body movements and emotions using fuzzy comprehensive evaluation. The two types of labels initially selected are evaluated by 20 experts in turn, giving the correlation scores. Combined with the previous datasets and evaluation scores, we determined the one-to-one correspondence between emotions and body movements, and established the mapping relationship, as shown in Figure 3.

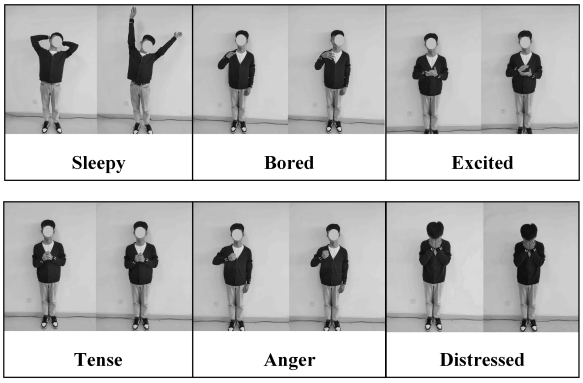

Figure 3 Illustration of the mapping relationship between emotions and body movements

3.2 Feature Extraction

Due to the large amount of emotional data collected by inertial sensors, it is difficult to obtain as much complete information as possible about emotions. Refer to [27], data segmentation technology based on sliding window is used, which includes two parameters: window width and overlap rate. The overlapping of adjacent windows is to avoid data loss at the edge of the window. The data is divided according to the overlap rate of 50% and the width of windows of 2 seconds (100 sampling points). Within this frame, a total of 6496 samples are obtained. Besides, the original data of the accelerometer, gyroscope, and magnetometer are fused through the Kalman filter algorithm [28].

We manually extract 6 features from the data segments in each window, including time domain features and frequency domain features. The time domain features are mean, variance, kurtosis, and correlation coefficient, the frequency-domain features are energy and magnitude. It constitutes a 216-dimensional emotional feature space (6 features $\times$ 3 axes directions $\times$ 3 IMUs $\times$ 4 sensors). Take the acceleration signal as an example, the formulas of features are as follows:

Mean:

$$\overline{a_x} = \frac{1}{\omega}\sum_{i=1}^{\omega} a_{x_i} \tag{1}$$

Variance:

$$\sigma(a_x) = \frac{1}{\omega}\sum_{i=1}^{\omega} (a_{x_i} - \overline{a_x})^2 \tag{2}$$

Kurtosis:

$$K = \frac{E(a_x - \overline{a_x})^4}{\sigma(a_x)^2} \tag{3}$$

Correlation coefficient:

$$R\left(a_x, a_y\right) = \frac{COV(a_x, a_y)}{\sqrt{\sigma(a_x)\sigma(a_y)}} \tag{4}$$

Energy:

$$E = \sum_{i=1}^{\omega} \left|d_{x_i}^2\right| \tag{5}$$

Magnitude:

$$Mag = \max(2 * |d_{x_i}|) \tag{6}$$

where ω is the width of the sliding window, a_x is the data of x-axis acceleration, a_y is the data of y-axis acceleration, a_{x_i} is the i-th data in a_x, $i = 1, 2, \cdots, \omega$, $COV(a_x, a_y)$ is the covariance of a_x and a_y, d_{x_i} is the FFT coefficient of a_{x_i}.

3.3 Feature Selection

There may be correlation and redundancy between the extracted emotional features, so it is necessary to simplify the data structure and improve the operation speed of the recognition algorithm. Principal component analysis (PCA) is to recombine the original features into linear independent new features. We select features with cumulative contribution rate higher than 95% and obtain 6 new features.

In order to obtain the original features that contribute greatly to the new features, the factor loadings between the new features and the original features are obtained through factor analysis techniques. Factor loading greater than 0.75 can be used to represent the original features that determine the main influence of the new features. It can be obtained that the mean, kurtosis, and energy contribute the most to the new features.

Since the unit and range of the emotional data collected by the accelerometer, gyroscope and magnetometer are inconsistent. Z-score [29] is used to standardize the input data to the classifier, the following is the formula:

$$X^* = \frac{X - \overline{X}}{s(X)} \tag{7}$$

where $\overline{X}$ is the mean, and $s(X)$ is the standard deviation.

3.4 Weighted Kernel-SVM Model

This paper aims to improve the accuracy of emotion recognition, proposes weighted kernel SVM model based on the SVM. A new kernel function is obtained by linear convex combination of multiple kernel functions, in which the weight of each kernel function is determined by the fuzzy function. Thus reducing the influence of the selection of a single kernel function on the recognition accuracy.

SVM is a supervised machine learning model that finds the best hyperplane to distinguish between positive and negative training samples. It can map the linearly inseparable problem to the high-dimensional space through the kernel function and convert it into linearly separable, so as to obtain the optimal solution. For the training sample $\{(\mathrm{x}_1, y_1), (\mathrm{x}_2, y_2), \cdots, (\mathrm{x}_N, y_N)\}, x_i \in R^n, y_i \in \{-1, +1\}, i = 1, 2, \cdots, N$ can be separated by the hyperplane $y_i\left(\omega^T x + b\right) \geq 1, i = 1, \cdots, n$, the SVM model can be expressed as

$$\min_{\omega, b, \xi} \frac{1}{2}\|w\|^2 + C\sum_{\mathrm{i}=1}^{N} \xi_i \tag{8}$$

$$\begin{aligned} s.t. \quad & \mathrm{y}_i\left[\omega\varphi\left(x_i\right) + b\right] \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \cdots, \mathrm{N} \end{aligned} \tag{9}$$

where $\varphi\left(x_i\right)$ is the feature map of vector x_i, which is the mapping function, $C > 0$ is the penalty parameter, ξ_i is the slack variable, b is the offset value, and ω is the normal vector of the separation hyperplane.

Compared with the single kernel SVM, the multi-core model has higher flexibility. The high-dimensional space after multiple kernel function mappings is a combination space formed by combining multiple feature spaces. Different feature components in heterogeneous data can be mapped through the most suitable single kernel function, finally, the data can be expressed more accurately and reasonably in the new combination space, thereby improving the prediction accuracy of the sample data. In general, it is a linear convex combination of multiple kernel functions, as follows

$$K\left(x, z\right) = a_1 K_1\left(x, z\right) + \alpha_2 K_2\left(x, z\right) + \cdots + \alpha_k K_k\left(x, z\right) \tag{10}$$

$$a_1 + \alpha_2 + \cdots + \alpha_k = 1 \tag{11}$$

where $K_k(x,z)$ is the predetermined kernel functions, α_k is the weight value, k is the number of kernel functions. $K(x,z)$ is the weighted kernel function, it also satisfies the Mercer theorem, the kernel function is defined as

$$K(x,z) = \varphi(x)\varphi(z) \tag{12}$$

where, $\varphi(x)\varphi(z)$ is the dot product of two.

In this paper, the fuzzy function is used as an auxiliary to obtain the subordinative degree of the model recognition accuracy, and then the accuracy of the single kernel SVM model is used to calculate the weight of the corresponding kernel function. The fuzzy function form is as follows, and its function image is shown in Figure 4.

$$\theta(\lambda) = \begin{cases} 0 & , \quad 0 < \lambda < a \\ \left(\frac{\lambda - a}{1-a}\right)^t & , \quad a \leq \lambda \end{cases} \tag{13}$$

where, $\theta(\lambda)$ is the membership function, λ is the accuracy of the model, t is the number of sensors, and a is the constant.

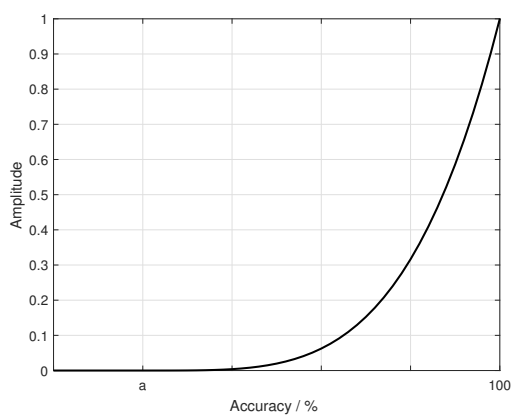

Figure 4 K-th semiparabolic distribution

Here, we believe that the recognition effect of SVM model is poor when the accuracy is lower than a. When the accuracy of the model is greater than a, the classification effect of the model meets the requirements, and the kernel function used by the model is suitable to be used as a member of the weighted kernel function. The value of a is not certain, but it could be assigned according to experience. In general, the greater the number of sensors, the better the recognition achieved. $\theta(\lambda)$ is the concave function between a and 1. As t increases, the discrimination of parts with relatively high recognition rate becomes more obvious. In this paper, the parameters $a = 0.6$ and $t = 4$ in the fuzzy function are set.

Due to the relatively large amount of emotional data, it takes a long time to use all the dataset to select the kernel functions, so we choose 30% of the data for prescreening. First, we use the commonly used single-kernel SVM to identify emotions, with kernel functions such as linear kernel function, polynomial kernel function,

radial basis function and Gaussian kernel function, and obtain the accuracy λ respectively. Then determine the degree of membership $\theta(\lambda)$ corresponding to each recognition accuracy according to the fuzzy function, and the classification weight α_i is calculated. The functional form of the classification weight is

$$\alpha_i = \frac{\theta(\lambda_i)}{\sum_{i=1}^{k} \theta(\lambda_i)}, i = 1, 2, \cdots, k \tag{14}$$

where k is the number of kernel functions, and $\theta(\lambda_i)$ is the subordinative degree corresponding to the accuracy of the model using the i-th kernel function.

4 System Platform and Data Collection

4.1 System Platform

This paper constructs a multi-sensor wireless transmission system platform composed of four wireless nine-axis inertial sensors, four Bluetooth signal receivers, and a notebook computer. Each sensor(WIT Inc., CHN) has built-in ICM42605(accelerometer, gyroscope) and MMC3630(magnetometer), which have the characteristics of small size, wearable, and low power consumption. Table 1 lists the concrete parameters.

Table 1 Sensor parameters

	Range	Error
Accelerometer	±16 g	0.01 g
Gyroscope	±35 rad/s	0.00087 rad/s
Magnetometer	±30 Gauss	2 mG

Each inertial sensors read nine analog-to-digital conversion output values, including three-axis acceleration, three-axis angular rate, and three-axis magnetic field. Then send the read data to the Bluetooth receivers(HC06-HID) through the Bluetooth protocol. HC06-HIDs transmit the data to the computer through USB to receive and process the nine-axis data, the whole process is shown in Figure 5.

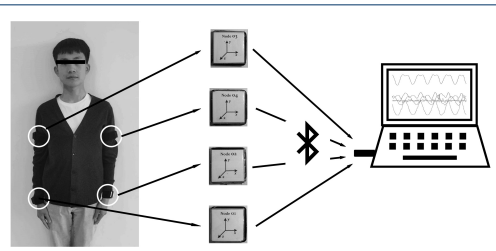

Figure 5 Construction of the system platform

4.2 Data Collection

A total of nine volunteers are selected for this paper to participate in the data collection, including five males and four females, aged from 22 to 26 years old, in

good health. Four sensors are bound to the left wrist, right wrist, left forearm, and right forearm of each volunteer. Set the sensor sampling frequency to 50 Hz, and collect each emotion data continuously for two minutes. During the process, the volunteers continue to repeat the actions according to their own habits and methods, but the normal data collection cannot be disturbed. Figure 6 shows the emotional data collected by the x-axis accelerometer in the sensor on the right wrist of the first volunteer in six emotions.

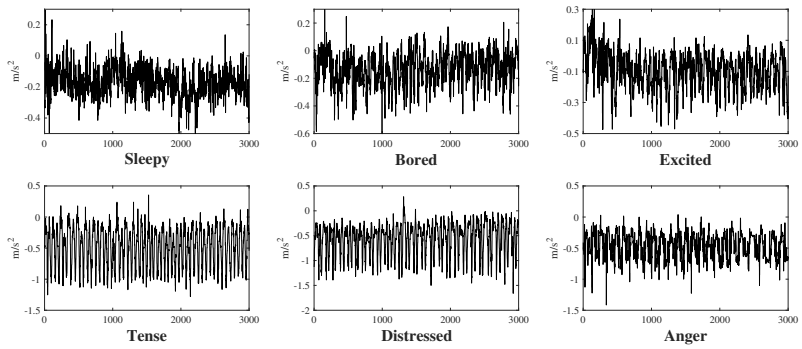

Figure 6 Data waveform of the x-axis accelerometer in the sensor on right wrist in six emotions

5 Experimental Results and Analysis

5.1 Parameter Optimization

The weight kernel SVM implementation is based on a LibSVM with one-vs-all strategy, and the regularization term C and G(in a range from 1×10^{-10} to 1×10^{10} logarithmically spaced). A grid search method is used to optimize the parameters by using Leave-One-Out cross-validation. 4000 samples are randomly selected as the benchmark dataset, 80% of the data are used as the training set, and the other 20% of the data are used as the verification set of the classification model. The process is repeated 5 times until each data is used as a validation set. The average accuracy of 5 iterations is used to evaluate the performance of the classifier. According to the highest accuracies, the weight kernel-SVM's regularization term C is set to 1.14 and G is set to 3.

5.2 Result

When the training model is completed, to determine the quality of the classifiers, the trained model needs to be evaluated. We use the following common evaluation indexes:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$Recall = \frac{TP}{TP + FN} \tag{17}$$

$$F_1 score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{18}$$

where TP is the number of correctly classified positive examples, FP is the number of positive examples misclassified as negative examples, TN is the number of correctly classified negative examples, FN is the number of negative examples misclassified as positive examples number.

Randomly select 1000 sample points in addition to the benchmark dataset as the test set to identify six emotions. Table 2 shows the precision, recall and F_1score of the six emotions under the weighted kernel SVM model. It can be seen from the results that the precision of tense and anger stand out, the recall of sleepy is the highest, and the F_1score of tense is outstanding. The confusion matrix in Figure 7 more clearly shows the six emotional recognition results, in which the vertical axis is the true value and the horizontal axis is the predicted value. The total number of data in each row represents the number of data instances of the category, and the value in each column represents the number of true data predicted to be the category.

Table 2 Precision, recall and F_1score of six emotions using weighted kernel-SVM model

No.	Emotion	Precision	Recall	F_1score (%)
1	Sleepy	96.53	**100**	98.23
2	Bored	97.67	98.25	97.96
3	Excited	98.31	97.75	98.03
4	Tense	**100**	99.42	**99.71**
5	Anger	**100**	98.00	98.99
6	Distressed	98.11	96.89	97.50

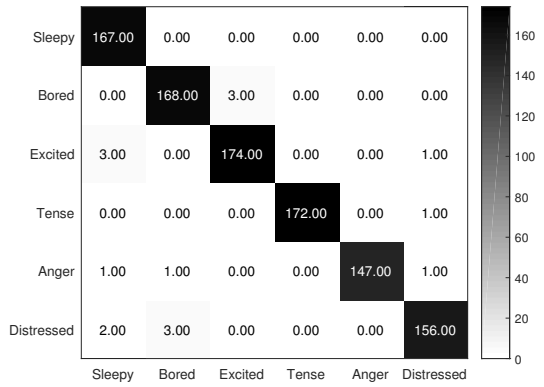

Figure 7 Confusion matrix of six emotions using weighted kernel-SVM model

5.3 Sensor System Minimization

To realize the possibility of changing the combination of sensors based on affecting the recognition effect as little as possible, the following six sensor combinations are selected to cooperate with weighted kernel SVM for simulation: accelerometer and gyroscope (A+G), accelerometer and magnetometer (A+M), gyroscope and magnetometer (G+M), accelerometer (A), gyroscope (G), Magnetometer (M).

The comparison of sensor combinations adopts the permutation importance method [30], and the data used is the feature data before PCA. We use all the sensor data for training, get a model, and use the validation set to get the standard score. Simply, multiply the weighted kernel SVM model accuracy by 100, using 98.4 as the standard score. Then randomly shuffling the features of the sensor combination we selected in the verification set for verification. Repeat the above process for all sensor combinations. The more the sensor combination score is reduced, the greater the importance of the sensor combination. In the simplest case, the sensor combination score is the recognition accuracy obtained after the sensor combination is shuffled, and the contribution score is the standard score minus the sensor combination score.

Figure 8 shows the accuracy and contribution of different inertial sensor combinations to the six emotion recognition effects. Among them, the accelerometer and magnetometer contributes the most to the accuracy, followed by the gyroscope and magnetometer. In contrast, when only the magnetometer is used, the contribution to the recognition effect is unsatisfactory. Therefore, it is recommended to use the combination of accelerometer and magnetometer or magnetometer to collect emotional data, this can reduce the number of sensors and maintain high accuracy.

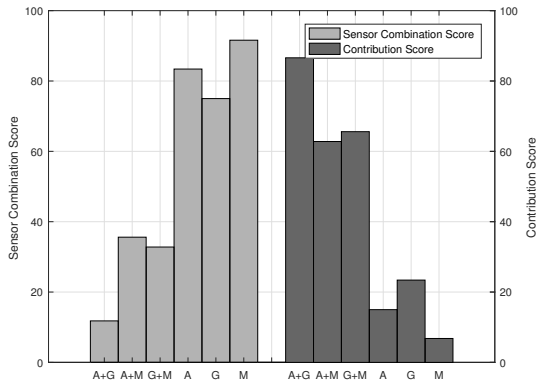

Figure 8 Comparison of socres between six different types of sensor combinations

6 Discussion

In this work, we use wearable inertial sensors for emotion recognition by establishing a mapping relationship between body movements and emotions. Research shows that human emotions can be recognized using inertial sensors, which is consistent with previous research[21, 22, 23].

We used four inertial sensors to fix the volunteers' wrists and arms, and collected six emotional data: sleepy, bored, excited, tense, anger and distressed. The data is segmented through a sliding window and features are extracted as the input of the classifier. The calculation method of the weights of the kernel functions is determined by the fuzzy function, and the weighted kernel-SVM is constructed. The

results show that the classification accuracy for the six emotions is 98.40% using weighted kernel-SVM. Table 3 shows the comparison of the accuracy of the sentiment classification model between the proposed system and other related works. Compared with RBF kernel-SVM, for six emotions, the accurary of weighted kernel-SVM is increased by 4.09%, this shows that the proposed method is effective.

Table 3 Performance comparison of other related systems

Reference	Classifier(s)	Accuracy	Precision	Recall	F_1score (%)
Cui et al. [31]	RBF-SVM	94.31	95.20	94.42	94.81
	Random Tree	89.80	88.46	91.54	89.97
	Multi-Layer Perception	92.11	92.48	91.95	92.21
	Decision Tree	88.56	90.58	86.07	88.27
	Random Forest	91.54	88.13	96.02	91.91
M.A. Hashmi et al. [21]	RBF-SVM	94.31	95.20	94.42	94.81
	Random Forest	91.54	88.13	96.02	91.91
Raffaele Gravina et al.[22]	HMM	76.24	78.02	77.35	77.68
Zhang et al.[23]	Random Tree	89.80	88.46	91.54	89.97
	RBF-SVM	94.31	95.20	94.42	94.81
	Random Forest	91.54	88.13	96.02	91.91
	Decision Tree	88.56	90.58	86.07	88.27
Proposed Approach	Weighted kernel-SVM	**98.40**	**98.44**	**98.39**	**98.41**

It is worth noting that the limitations of this study need to be pointed out. This study recruited 9 volunteers to collect emotional data, and the number of sample sets was limited. We can increase the number of volunteers and increase the dataset capacity. The data in this study can be tested as a test set to verify the effectiveness of the proposed method. In addition, this paper uses the traditional feature extraction method. The manually selected features are intuitive, but it takes time and effort in feature design. So we consider designing an automatic feature extraction module using deep neural network. In order to obtain better effect of emotion classification, research based on multi-modality is becoming more and more extensive. it is planned to fuse sensor data and images, inertial signals and physiological signals. These limitations point out the direction for future work.

7 Conclusion

This paper establishes the mapping relationship between emotion and movement, and uses inertial sensors to collect data. Comparing the experimental results, the emotion recognition method based on wearable inertial sensors and weighted kernel-SVM can be applied to certain scenes in life, which can bring convenience to people.

Declarations

Abbreviations

SVM: support vector machine; IMU: inertial measurement units; FNN: fuzzy neural network; CNN: convolutional neural network; 1D: one-dimensional; 2D: two-dimensional; PCA: principal component analysis.

Acknowledgements

The authors would like to thank the volunteers who provided service and assistance in the data collection efforts.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

All authors read and approved the fnal manuscript.

Funding
This work was supported by National Natural Science Foundation of China under (Grant Nos. , 61903170, 62173175, 61877033), and by the Natural Science Foundation of Shandong Province under grants Nos. ZR2019BF045, ZR2019MF021, and by the Key Research and Development Project of Shandong Province of China, No.2019GGX101003.

Availability of data and materials
Please contact the authors for data requests.

Ethics approval and consent to participate
This article is ethical, and this research has been agreed.

Consent for publication
The picture materials quoted in this article have no copyright requirements, and the source has been indicated.

References
1. Koonin, E.V., Altschul, S.F., Bork, P.: Brca1 protein products: functional motifs. Nat Genet **13**(2), 266–267 (1996)
2. Yan M C, O.: Statistical recovery of wideband speech from narrowband speech. IEEE Trans Speech & Audio Processing **2**(4), 544–548 (1994)
3. Ming Guo, N.Y. Zhelong Wang: A multisensor multiclassifier hierarchical fusion model based on entropy weight for human activity recognition using wearable inertial sensors. IEEE Transactions on Human-Machine Systems **49**(1), 105–111 (2019)
4. Zhelong Wang, C.Z. Ming Guo: Badminton stroke recognition based on body sensor networks. IEEE Transactions on Human-Machine Systems **46**(5), 769–775 (2016)
5. Meeren H K, D.G.B. Heijnsbergen C V: Rapid perceptual integration of facial expression and emotional body language. Proceedings of the National Academy of Sciences of the United States of America **102**(45), 16518–16523 (2005)
6. Saurav S, S.S. Saini R: Emnet: a deep integrated convolutional neural network for facial emotion recognition in the wild. Applied Intelligence **51**(8), 5543–5570 (2021)
7. Kulkarni P, R.T.M.: Video based sub-categorized facial emotion detection using lbp and edge computing. Revue d'Intelligence Artificielle **35**(1), 55–61 (2021)
8. Sanfilippo F, P.K.Y.: A sensor fusion wearable health-monitoring system with haptic feedback. In: 2015 11th International Conference on Innovations in Information Technology (IIT), pp. 262–266 (2015)
9. Akh A, S.B.: A novel facial emotion recognition scheme based on graph mining. Defence Technology **16**(5), 1062–1072 (2020)
10. Kaviya P, A.T.: Group facial emotion analysis system using convolutional neural network. In: 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), pp. 643–647 (2020)
11. Nicolai A, C.A.: Facial emotion recognition using fuzzy systems. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2216–2221 (2015)
12. Shi W, J.M.: Fuzzy wavelet network with feature fusion and lm algorithm for facial emotion recognition. In: 2018 IEEE International Conference of Safety Produce Informatization (IICSPI), pp. 582–586 (2018)
13. Jian-Hua Tao, Y.L.Z.L.M.-Y.N. Jian Huang: Semi-supervised ladder networks for speech emotion recognition. International Journal of Automation and Computing **16**(4), 437–448 (2019)
14. Jiang Xiaoqing, L.Y.B.J. Xia Kewen: Noisy speech emotion recognition using sample reconstruction and multiple-kernel learning. Journal of China Universities of Posts & Telecommunications **24**(2), 1–9 (2017)
15. Jianfeng Zhao, L.C. Xia Mao: Learning deep features to recognise speech emotion using merged deep cnn. IET Signal Processing **12**(6), 713–721 (2018)
16. Shaoling Jing, L.C. Xia Mao: Prominence features: Effective emotional features for speech emotion recognition. Digital Signal Processing **72**, 216–231 (2018)
17. Yuan Wu, Q.D.S.T. Haibing Zhu: A survey of the research status of pedestrian dead reckoning systems based on inertial sensors. International Journal of Automation and Computing **16**(1), 65–83 (2019)
18. Ming Guo, N.Y. Zhelong Wang: A multisensor multiclassifier hierarchical fusion model based on entropy weight for human activity recognition using wearable inertial sensors. IEEE Transactions on Human-Machine Systems **49**(1), 105–111 (2019)
19. Z Wang, Z.C.J.M. S Qiu: Quantitative assessment of dual gait analysis based on inertial sensors with body sensor network. Sensor Review **33**(1), 48–56 (2013)
20. Motti Ader Lilian Genaro, M.K.C.B. Greene Barry R.: Reliability of inertial sensor based spatiotemporal gait parameters for short walking bouts in community dwelling older adults. Gait & Posture **85**, 1–6 (2021)
21. M. A. Hashmi, M.Z.M.S.M.M.F. Q. Riaz: Motion reveal emotions: Identifying emotions from human walk using chest mounted smartphone. IEEE Sensors Journal **20**(22), 13511–13522 (2020)
22. Rg A, Q.B.: Emotion-relevant activity recognition based on smart cushion using multi-sensor fusion. Information Fusion **48**, 1–10 (2019)
23. Zhan Zhang, L.C.X.L.T.Z. Yufei Song: Emotion recognition based on customized smart bracelet with built-in accelerometer. Peerj
24. A, R.J.: A circumplex model of affect. Journal of Personality and Social Psychology **39**(6), 1161–1178 (1980)
25. Scherer K R, E.H.: Multimodal expression of emotion: affect programs or componential appraisal patterns? Emotion **7**(1), 158–71 (2007)
26. Noroozi, F., Corneanu, C., Kaminska, D., Sapinski, T., Escalera, S., Anbarjafari, G.: Survey on emotional body gesture recognition. IEEE Transactions on Affective Computing **12**(02), 505–523 (2021)
27. Attila Reiss, D.S.: Aerobic activity monitoring: towards a long-term approach. Universal Access in the Information Society **13**(1), 101–144 (2014)

28. Wang X M, N.W.B. Liu Z: Motion measurement and recognition method of human arm based on mems sensors. Journal of Chinese Inertial Technology **25**(6), 701–707 (2017)
29. V, J.: Computational statistics handbook with matlab. Journal of the Royal Statistical Society **172**(4), 942–943 (2009)
30. A. Calado, G.S. V. Errico: Toward the minimum number of wearables to recognize signer-independent italian sign language with machine-learning algorithms. IEEE Transactions on Instrumentation and Measurement **70**, 1–9 (2021)
31. Cui L, Z.T. Li S: Emotion detection from natural walking. In: Human Centered Computing, pp. 23–33 (2016)