



Wearable IMU-based Gym Exercise Recognition Using Data Fusion Methods

Jiacheng Tian
University of science and technology
of China Anhui, China
tien1012@outlook.com

Panzhou
Shenzhen Institute of Advanced
Technology, Chinese Academy of
Sciences, Shenzhen, China
pan.zhou@siat.ac.cn

Fangmin Sun*
Shenzhen Institute of Advanced
Technology, Chinese Academy of
Sciences, Shenzhen, China
fm.sun@siat.ac.cn

Tao Wang
China Astronaut Research and
Training Center, Beijing China
wangjiyuanwpt@126.com

Hexiang Zhang
China Astronaut Research and
Training Center, Beijing China
604391781@qq.com

ABSTRACT

Gym exercise has become a focus of attention nowadays because of its health benefits. Automatic gym exercise recognition is an emerging research field which aimed at guiding people to keep fit scientifically through gym exercise monitoring. However, as the actions of gym exercise (e.g. barbell bench press, leg extension, etc.) are more diversity and complexity than outdoor exercise (e.g. running, cycling, etc.). Previous studies increase the number of sensors to improve the accuracy gym exercise recognition, while wearing too many sensors make the subjects uncomfortable during gym exercise. In this study, we studied the performance of different classifiers on gym exercise recognition, then the impact of the number of sensors and the positions of the sensor on the body on the recognition performance was analyzed based on the Extra Trees (ET) classifier. Finally, a stratification fusion method using only two sensors was proposed according to the analysis results. The experimental results showed that when two sensors were used to identify eight kinds of gym exercises, the accuracy of the proposed stratification fusion method was 91.26%.

CCS CONCEPTS

• Computing methodologies; • Machine learning;; • Artificial Intelligence;;

KEYWORDS

Gym exercise recognition, Multiple wearable sensors, Stratification fusion methods

ACM Reference Format:

Jiacheng Tian, Panzhou, Fangmin Sun*, Tao Wang, and Hexiang Zhang. 2021. Wearable IMU-based Gym Exercise Recognition Using Data Fusion Methods. In *The Fifth International Conference on Biological Information and*

Biomedical Engineering (BIBE2021), July 20–22, 2021, Hangzhou, China. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3469678.3469705>

1 INTRODUCTION

As a subset of physical activity, gym exercise is a planned and periodically activity that requires professional guidance and aims at promoting physical quality [1]. A better effect of exercise can be obtained under a scientific exercise guidance, which has significantly impact on staying healthy and rehabilitating from diseases. Consequently, an effective plan and accurate records of the exercise are necessary. However, the recording method is mainly manual, which is time consuming, laborious, and reliant on the instructor's professionalism.

Researches on human activity recognition (HAR) based on wearable sensors, such as accelerometers, gyroscopes, and global positioning system (GPS), has been studied for many years. The earliest wearable sensors based HAR can be traced back to the study by Van Laerhoven in 2000 [2] in which daily activity recognition was first achieved using two sensors installed in trousers. Nowadays, daily activity recognition has been widely studied, several studies [3–5] have achieved acceptable performance on daily activity recognition using different methods. However, the daily activities recognized by these studies including walking, running, lying, etc. are the coordinated activity of the whole body, which can be recognized by the information sampled by a single sensor.

Due to the complexity and diversity of fitness actions in the gym, the corresponding recognition method still have limitations at certain cases. In the last decade, with the popularity of inertial sensor units, it is common for people to wear multiple inertial sensor units, which make it possible to use wearable inertial sensors for automatic gym exercise recognition [6]. Those wearable inertial sensors that integrate accelerometers, gyroscopes, magnetometers and other elements can usually be called Inertial Measurement Units (IMU).

Recently, gym exercise recognition, as a part of HAR, has begun to attract the attention of researchers [4, 7]. According to the purpose of gym exercise recognition, previous researches are generally fall into two groups. The first group was the fine-grained recognition of specific gym exercises which mainly aimed at revealing the relationship between limb motion and muscle. However, these methods usually require the users to wear multiple sensors on a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

BIBE2021, July 20–22, 2021, Hangzhou, China

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8929-7/21/07...\$15.00

<https://doi.org/10.1145/3469678.3469705>

specific body part which make the user uncomfortable. The other group mainly focuses on the coarse-grained recognition of gym exercises which aimed at recognizing the type of exercise the user was doing. Although this group of methods needs fewer sensors, the recognition accuracy is very low.

Zhang proposed a Fine-fit system that used sensor networks to recognize fine-grained level of gym exercises, including push-ups and barbell curl, and its prediction accuracy was about 91% [8]. Some studies used inertial sensor signals to recognize different kinds of activities [9, 10], while others recognized activities by combining acceleration data and heart-rate monitoring signal [11]. Heli, et al used a single accelerometer to recognize 30 different upper-limb gym exercises with accuracy reaching 96% [12]. However, previous studies on coarse-grained gym exercise recognition had certain limitations. In the case of using one or two sensors, better accuracy was obtained in the recognition of human upper limb (or lower limb). However, recognizing the whole-body movement (including upper limb and lower limb) was impossible, unless the number of wearable sensor nodes was increased.

Multi-sensor information fusion method refers to fusing the information collected from multiple sensors to achieve a better performance. However, how to optimize the number of sensors and the performance in different application scenarios is still a problem to be explored. In this paper, firstly, we analyzed the impact of the number of sensors and different classifiers on the recognition performance. Then, based on the analysis results of the impact, we proposed a data fusion method, which is multi-sensor stratification fusion method to realize the recognition of human gym exercises. And the main contributions of this paper are as follows:

1. We analyzed the impact of the number and the position of sensors and the types of classifiers on the recognition performance. In order to analyze the impact of the number and the position of sensors and the types of classifiers on the recognition performance, fusion methods were used in two experiments. (1) To analyzing the impact of classifier on recognition performance, we compared the performance of four single classifiers and the fusion method. The gym exercise recognition accuracy is 95.02% when the entropy voting method is used. (2) Analyzed the impact of the sensor numbers and sensor positions on the recognition performance. Based on the entropy voting fusion method, the influence of different positions on the recognition performance was analyzed by changing the number and position of sensor nodes. The experimental results showed that the nodes with hand and ankle positions account for a large weight in the recognition. However, the accuracy of gym exercise recognition was only 85.2% when two sensors (hand and ankle) were used.

2. We proposed a stratification based fusion method to improve the gym exercise recognition performance when using only two sensors. The gym exercises were divided into the upper and lower limbs, the upper layer and lower layer models were separately trained using a single classifier, then, the entropy voting method was used to get the final prediction. The experimental results showed that an accuracy of 91.26% could be obtained when wearing two sensors (hand and ankle).

The following of this paper is organized as follows: We introduced the experiments in Section II; in Section III we verified the



Figure 1: Data Collection Scenario.

effectiveness of the fusion method; finally, the conclusions were drawn in Section IV.

2 EXPERIMENTS

2.1 Data Collection

We used the Shimmer3 IMU wireless sensor and ConsensysPRO software produced by Shimmer for data collection. 6DoF inertial data including 3-axis acceleration and 3-axis gyroscope data were collected with a sample frequency of 100Hz.

A total of 12 subjects whose ages ranges from 22 to 27 were invited to participate in the experiment. The volunteers were informed of the purpose of the experiment, the steps, the way the nodes were worn and the recording rules of the action tags before experiment. Five sensors were placed on the left ankle, left leg, waist, left wrist and left arm of the subjects separately using the buckle. Each volunteer was asked to do 8 kinds of gym exercise (Pec fly, Running, Bicycle, Barbell bench press, Leg extension, Prone leg curl, Pull down, Sit-ups), which include 4 upper limb exercises and 4 lower limb exercises, and each kind of gym exercise includes 10 groups. An example of the sensors' locations are shown in Figure 1

As shown in Figure 1, this study collects data in the gym, and the experiment does not have uniform requirements for volunteers, including equipment weight-bearing, movement sequence and action standards. Finally, a total of 4800 gym data samples were collected (12 people * 8 movements * 10 groups * 5 sensors), and for each sample, the 6-axis information of acceleration and angular velocity is included.

2.2 Data Processing

Data preprocessing mainly included three procedures namely data filtering, data segmentation, and feature extraction. The noise was mainly originated from: (1) The inherent noise of the hardware of the inertial sensor; (2) the slight sensor position deviation and the noise caused by jitter in the process of exercise. The results of the experiment revealed that the human motion frequency was generally less than 20 Hz. Thus, the purpose of the noise filtering was mainly aimed at suppressing the high-frequency jitter signal. In this study, a fourth-order moving average filtering algorithm was used to suppress the high-frequency jitter noise, as this algorithm used the average of continuous N sampling points to replace the

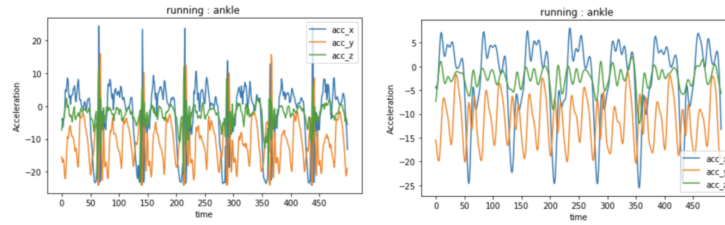


Figure 2: Comparison of Accelerometer Three-Axis Data Before and After During Filtering.

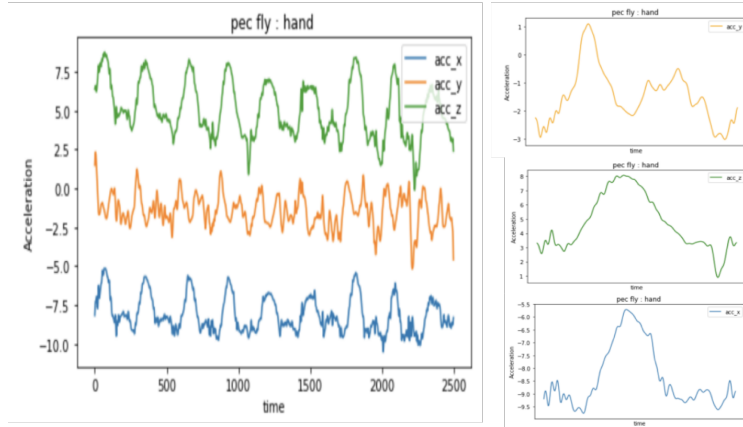


Figure 3: Detail of Accelerometer Three-Axis Data During Sliding Window.

sampling points, so it can make the signal smoother. An example of the filtering result was shown in Figure 2

The sensor data stream was segmented by an effective fixed-length sliding window segmentation method. The window length was 2s, and the adjacent windows had 50% overlap to avoid the loss of information. One segmentation result was shown in Figure 3. Eight features were extracted from each window, including the frequency and time domain features. The time domain feature mainly included the mean, variance, mean square deviation, skewness, kurtosis, and maximum and minimum of the signal (three-axis acceleration and angular velocity data), while the frequency domain feature included energy. Therefore, 48 dimensions (8 features \times 6 axes) were extracted for each sample feature.

The dataset of this study comprised the feature vectors of 12 volunteers. The independent feature vectors of 2 volunteers were taken as the test set and the feature vectors of the other 10 volunteers were merged as the train set to verify the performance of accuracy and generalization of the model. The cycle was repeated for 12 times, and the average was taken as the final result.

3 MODEL AND ANALYSIS

The HAR performance based on the wearable sensor is affected by many factors, including the performance of IMU sensors, the complexity of action, and the number of sensors, etc., while the most important factor is the performance of the classifiers. In the gym scene, the muscle group of limb exercises through equipment includes the upper and lower limbs. Although limb has

the characteristics of repetition and coherence during mechanical exercises, resisting gravity during the exercise period is required. Therefore, the data of inertial sensors having a mass of noise, and it is hard to obtain preferable hypothetical classification results using only one or two sensors. In traditional method, experiments were performed to determine the optimal single classifier. This method requires a prior information. If the prior information is insufficient, the model is not necessarily efficient.

Overall, in this Section, for the purpose of using fewer sensors and fewer classifiers to recognize gym exercise, we analyzed the impact factors step by step through three experiments. Firstly, Experiment 3.1 verified that the fusion method is more effective for gym exercise recognition. Then, on this basis, Experiment 3.2 studied the effect of sensor position and sensor number on the performance of gym exercise recognition. Finally, according to the results of Experiment 3.1 and 3.2, Experiment 3.3 proposed a stratification fusion method which used only 2 sensors to recognize gym exercise.

3.1 Multiple Sensor, Multiple Classifier Entropy Voting Method

Majority voting method: The majority voting algorithm is based on the law of large numbers and it has been widely used in many applications. It solves the conflict between multiple decisions and outputs the optimal decision of the system. The majority voting methods can be classified into three types: absolute majority voting, relative majority voting, and weight voting. For the absolute majority voting, the minority is subordinate to the majority. The

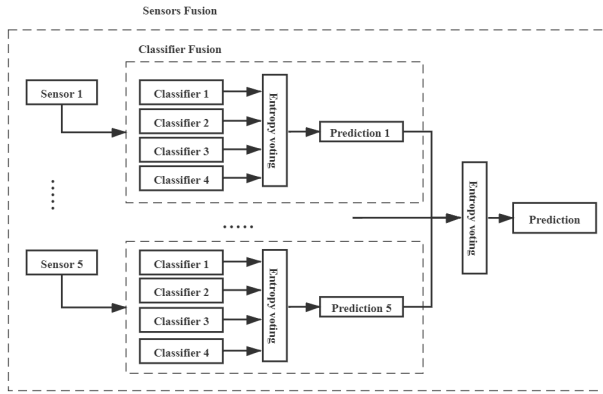


Figure 4: The Framework of the Multiple Sensors, Multiple Classifiers Method.

classification result is decided by the prediction labels which get the highest number of votes. For the relative majority voting, the classification result is outputted if it has more than a half votes. The weight voting algorithm based on the hypothesis that the performance of different classifier is diverse. The classifiers are given different weights based on their performance, the final prediction is the sum of the weighted prediction of each classifier, and the output label would be the label that got the maximum value. In this study, the entropy voting method was used in the multiple sensors, multiple classifiers model. The model of the proposed method is shown Figure 4

As shown in the Figure 4, the first layer of the model is the classifier fusion layer, which integrates four different single classifiers. Then the prediction results of the corresponding sensors are obtained by giving different weights to the classifier according to the entropy method. The second layer is the sensor fusion layer, which fuses the prediction results obtained by the five sensors in the first layer to get the final prediction result and the weights of different sensors are based on entropy.

The performance of the entropy voting method is verified by two experiments. The results are shown in Figure 5. In this experiment, the single classifier method is compared with their corresponding voting method. In Figure 5, the vertical coordinates represent the recognition accuracy and the horizontal coordinates represent the classifier category.

Figure 5 (left) the blue column represents the recognition accuracy directly obtained by a single classifier, and the green column represents the recognition accuracy obtained by combining the corresponding single classifier method with the proposed voting method. The blue bar chart in Figure 5 (right) shows a voting fusion algorithm that integrates the four single classifiers mentioned earlier, while the green bar chart adds an entropy value on a secondary basis.

As shown in Figure 5, a higher recognition accuracy could be achieved by the proposed fusion method when compared with single classifier. This indicated that the proposed fusion method was effective in human gym exercise recognition. A performance

Table 1: Comparison results of three methods

Method	Classifier	Accuracy%	Recall%	F1-score%
Single Classifier method	Random Forest	86.33	85.24	85.98
	Gradient Boosting	87.20	86.19	85.12
	Decision Tree	80.23	79.99	80.05
	Extra Tree	89.89	88.90	89.63
	Random Forest	89.89	88.90	89.63
Voting method	Gradient Boosting	88.50	88.36	88.03
	Decision Tree	84.58	82.28	83.40
	Extra Tree	91.01	90.89	90.89
	Voting	93.91	93.50	92.99
	Entropy Voting	95.10	94.99	95.02

comparison of different methods was shown in Table 1. The maximum accuracy achieved by the single classifier method was 89.98%. While the performance of fusion method improved by 1% to 5% compared with the single classifier method.

According to Table 1, the maximum accuracy achieved by the single classifier method was 89.98% which is much lower than the entropy voting based fusion method. While the performance of fusion method improved by 1% to 5% compared with single classifier method.

In addition, the accuracy of the entropy voting method was 95%, improved by 3% compared with the voting method without entropy. The conclusion drawn from the above experiments is that the entropy voting method has superior performance on gym exercise recognition when using 5 sensors. The confusion matrix was shown in Figure 6

3.2 Analysis of the Impact of the Number and Position of Sensors on Recognition

The fusion method introduced in 3.1 used five sensors, too many sensors make the subjects uncomfortable in practical scenarios. Therefore, in this section we compare the influence of the number of sensors on the recognition accuracy.

The test results are shown in Figure 7, the horizontal coordinates represent the combination of the number of different sensors, and the vertical coordinates represent the accuracy of the recognition classifier. The results shown that when five sensors are used, the highest accuracy is 95.10%, and if two sensors (hand + ankle) are selected, the accuracy is 86.31%. Although the recognition accuracy is higher when using five sensors, it is not suitable for real application. For the other hand, the recognition performance is inferior when using two sensors.

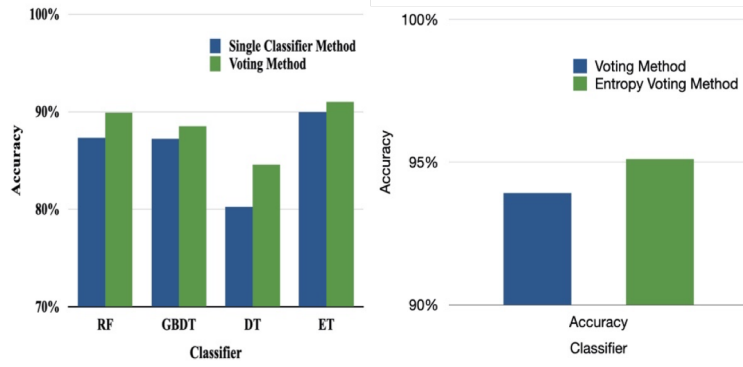


Figure 5: Comparison results of different method

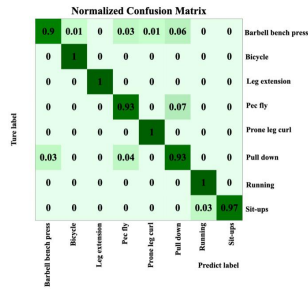


Figure 6: The Confusion Matrix of the Proposed Multi-Sensor and Multi-Classifier Entropy Voting Method.

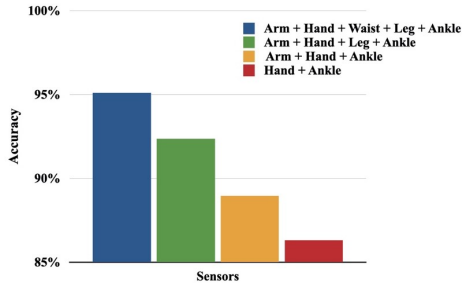


Figure 7: Accuracy Comparison When Using Different Number of Sensors.

3.3 Multiple Sensors Stratification Fusion Method

Gym exercises are usually local body movement, and the signals collected by sensors placed in different body parts will be very different. Figure 8 and 9 gave examples of the acceleration signal collected by two sensors located in ankle and hand when the subject doing pec fly and leg extension. It can be seen that during pec fly, the data of the sensor integrated on the ankle is relatively smooth, while the data of the hand sensor fluctuates greatly. However, during leg extension, the data collected by sensors located on ankle and hand is also very different.

Therefore, according to the characteristics of limb action during gym exercise, this paper proposes a stratification method to improve the gym exercise recognition rate while reducing the number of sensors. It is divided into upper and lower levels according to the types of exercises.

The **multiple sensor stratification fusion model** is shown in Figure 10. The basic layer contains a two-classifier with superior performance, through which sensor data can be divided into upper and lower layers as accurately as possible. These two layers contain the upper limb exercises label and the lower limb exercises label, namely upper label and lower label.

Upper layer and lower layer have their single classifiers respectively, which are used to further recognize specific body parts movements. By combining the upper prediction with the lower prediction, the final prediction is our original goal: recognize gym exercise.

The effectiveness of the method proposed in this paper is verified by experiments, and the experimental results are as follows: The basic layer performance affected directly the classification performance of the next layer. The exercise limbs needed to be classified correctly in this layer, and lower limbs and upper limbs needed to be divided separately into the upper layer and lower layer, respectively.

Table 2 shows the performance of metrics of different classifiers in the basic layer to reflect the basic classifier ability to stratify the limbs, namely upper layer and lower layer. As shown in Table 2, the basic layer had the best partition performance for limb exercises. The recognition accuracy reached 98.9%, and the recall and F1-score was 98.03%. So, the RF classifier was used in the following proposed methods.

Figure 11 shows the comparison of the recognition accuracy of the upper and the lower layers when using different single classifiers. The dark-green bar represents the performance of upper layer recognition during a gym exercise, and different classifiers represent the comparison of the recognition ability of the upper layer exercises. The blue bar represents the accuracy of lower layer recognition during gym exercise. The light-green bar represents the accuracy of the final recognition of human gym exercise using the proposed stratification fusion method. The recognition accuracy of all single classifiers in the upper layer was relatively lower than that of the lower layer, and the final recognition accuracy was higher than 90% when the ET classifier was used.

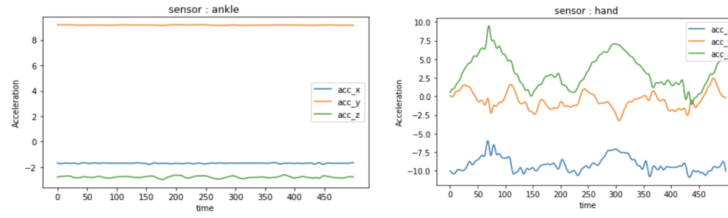


Figure 8: Three-axis Acceleration Sample During Pec Fly

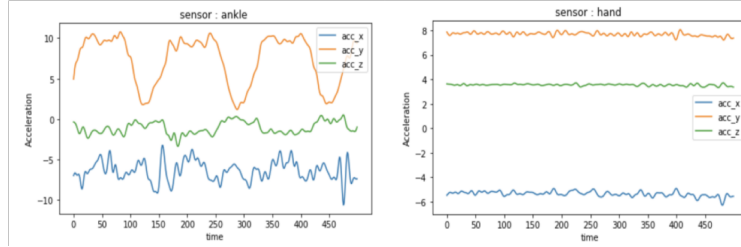


Figure 9: Three-axis Acceleration Sample During Leg Extension

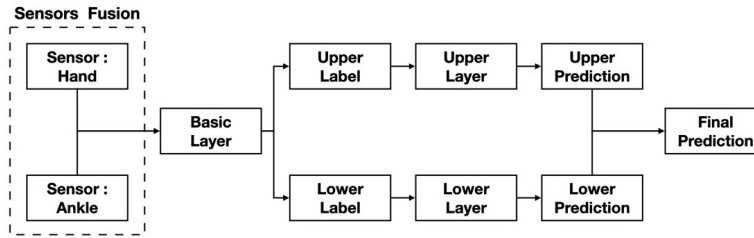


Figure 10: The Framework of the Proposed Stratification Fusion Method

Table 2: Performance Comparison Between DT Classifier and RF Classifier in the Basic Layer.

Classifier	Accuracy(%)	Recall(%)	F1-score(%)
DT	96.65	96.08	95.79
RF	98.90	97.89	98.03

The comparison between the multiple classifier, multiple sensor fusion method, and multiple sensor stratification fusion method in this experiment on gym exercise recognition in terms of different metrics on the basis of two sensors are shown in Table 3. The comparison shows that the recognition accuracy of the newly proposed method was better than that of the former, with a recognition accuracy of 91.26%; both recall and F1-score were more than 90%, which was obviously better than that using the previous method. Therefore, the methods combining two sensors to stratify and fuse have a practical value.

4 CONCLUSION

In this study, a data fusion method is proposed through three experiments. In this study, eight kinds of gym exercises are identified by

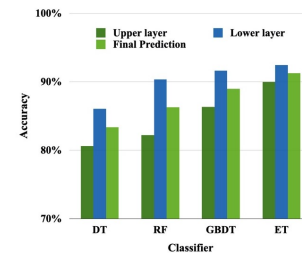


Figure 11: Performance Comparison of Four Types of Single Classifiers in Upper Layer, Lower Layer and Their Corresponding Final Predictions.

using wearable IMU sensors. An independent testing dataset is used to verify the effectiveness of the proposed recognition method. We analyzed the impact of the number and location of sensors and the types of classifiers on the recognition method. The results show that when the 2 sensors are combined, the accuracy can reach 86.31%. To reduce the number of sensors needed and improve the recognition accuracy, a stratification fusion method was then proposed. The fusion accuracy of the 2 sensors is 91.25%. The model is divided into upper and lower layers according to the features of gym exercises.

Table 3: Performance Comparison Between Entropy Voting Method and Stratification Voting Method When Using Two Sensors.

Method	Accuracy(%)	Recall(%)	F1-score(%)
Entropy Voting	86.31	85.23	85.98
Stratification Fusion	91.26	90.96	91.17

The experimental results show that the recognition accuracy of this method is superior to that of the single classifier methods using the same number of sensors.

ACKNOWLEDGMENTS

This study was sponsored by the National Natural Science Foundation of China with grant numbers 62073310 and 61902388; the basic research project of Guangdong Province with grant number 2021A1515011838; the foundation of National Key Laboratory of Human Factors Engineering with grant number 6142222180502; the National Key R&D Program of China with grant number 2018YFB1307005; Joint Fund of NSFC and Guangdong Province with grant number U1801261; Strategic Priority CAS Project with grant number XDB38040200; Basic research project of Shenzhen Science and Technology Innovation Commission with grant number JCYJ20180703145002040 and JCYJ20190807161805817.

REFERENCES

- [1] Zhang S, Li Z, Nie J, *et al.* (2016). How to record the amount of exercise automatically? A general real-time recognition and counting approach for repetitive activities. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2016, 831-834.
- [2] Van Laerhoven K, Cakmakci O. (2000). What shall we teach our pants? Digest of Papers. *Fourth International Symposium on Wearable Computers*. IEEE, 2000, 77-83.
- [3] Margarito J, Helaoui R, Bianchi A M, *et al.* (2015). User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach. *IEEE Transactions on Biomedical Engineering*, 63(4), 788-796.
- [4] Xiao F, Chen J, Xie X H, *et al.* (2018). SEARE: A system for exercise activity recognition and quality evaluation based on green sensing. *IEEE Transactions on Emerging Topics in Computing*, 2018.
- [5] Tapia E M, Intille S S, Haskell W, *et al.* (2007). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. *11th IEEE international symposium on wearable computers*. IEEE, 2007, 37-40.
- [6] Koskimäki H, Siirtola P. (2014). Recognizing gym exercises using acceleration data from wearable sensors. *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. IEEE, 2014, 321-328.
- [7] API,"<https://developer.android.com/reference/com/google/android/gms/location/package-summary.html>
- [8] Zhang Z, Wang N, Cui L. Fine-Fit. (2018). A Fine-grained Gym Exercises Recognition System. *24th Asia-Pacific Conference on Communications (APCC)*. IEEE, 2018, 492-497.
- [9] Koskimäki H, Siirtola P. (2014). Recognizing gym exercises using acceleration data from wearable sensors. *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*. IEEE, 2014, 321-328.
- [10] Zhang Z, Wang N, Cui L. Fine-Fit. (2018). A Fine-grained Gym Exercises Recognition System. *24th Asia-Pacific Conference on Communications (APCC)*. IEEE, 2018, 492-497.
- [11] Sun J, Li H. (2008). Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers. *Expert Systems with Applications*, 35(3), 818-827.
- [12] Khan A M, Lee Y K, Lee S Y, *et al.* (2010). A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE transactions on information technology in biomedicine*, 14(5), 1166-1172.