

A Novel Method for Cross-Subject Human Activity Recognition with Wearable Sensors

Qi Zhang¹, Feng Jiang², Xun Wang³, Jinnan Duan³, Xiulai Wang^{4*}, Ningling Ma⁴, Yutao Zhang⁴

¹Faculty of Computing, Harbin Institute of Technology, Harbin, China

²School of Future Technology, Nanjing University of Information Science & Technology, Nanjing, China

³School of Medicine and Health, Harbin Institute of Technology, Harbin, China

⁴Nanjing Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China

Email: *wangxiulai@126.com

How to cite this paper: Zhang, Q., Jiang, F., Wang, X., Duan, J.N., Wang, X.L., Ma, N.L. and Zhang, Y.T. (2024) A Novel Method for Cross-Subject Human Activity Recognition with Wearable Sensors. *Journal of Sensor Technology*, **14**, 17-34.

<https://doi.org/10.4236/jst.2024.142002>

Received: May 12, 2024

Accepted: June 25, 2024

Published: June 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Human Activity Recognition (HAR) is an important way for lower limb exoskeleton robots to implement human-computer collaboration with users. Most of the existing methods in this field focus on a simple scenario recognizing activities for specific users, which does not consider the individual differences among users and cannot adapt to new users. In order to improve the generalization ability of HAR model, this paper proposes a novel method that combines the theories in transfer learning and active learning to mitigate the cross-subject issue, so that it can enable lower limb exoskeleton robots being used in more complex scenarios. First, a neural network based on convolutional neural networks (CNN) is designed, which can extract temporal and spatial features from sensor signals collected from different parts of human body. It can recognize human activities with high accuracy after trained by labeled data. Second, in order to improve the cross-subject adaptation ability of the pre-trained model, we design a cross-subject HAR algorithm based on sparse interrogation and label propagation. Through leave-one-subject-out validation on two widely-used public datasets with existing methods, our method achieves average accuracies of 91.77% on DSAD and 80.97% on PAMAP2, respectively. The experimental results demonstrate the potential of implementing cross-subject HAR for lower limb exoskeleton robots.

Keywords

Human Activity Recognition, Cross-Subject Adaptation, Semi-Supervised Learning, Wearable Sensors

1. Introduction

With the rapid development of technologies such as robots, edge computing, and wearable devices, the exoskeleton robot has been extensively applied in civilian fields such as human augmentation, industrial production, and rehabilitation training, as well as in the military field [1]-[5]. As a wearable robot, an exoskeleton robot can enhance the endurance, load capacity, and mobility of the wearer while also reducing energy consumption. As a human-machine interaction system, exoskeleton robots need to satisfy urgent demands from users, such as accurately identifying and predicting wearer's movements in complex environments, supporting various types of movement modes, and achieving bidirectional adaptability between man and machine. Consequently, research gradually shifts to the upper control layer, *i.e.*, through accurate recognition and prediction of human movements via HAR, and then achieving bidirectional adaptability between man and machine through the cooperation of lower control strategies and the support of mechanical structures.

Wearable movement sensors are widely used in the field of HAR. Among various types of movement signals, acceleration signals are commonly used for human behavior recognition tasks, typically measured by inertial measurement units (IMUs) and providing the target's acceleration in three axes relative to the coordinate system. The advantages of this signal are high signal quality, but it cannot complete collection before the target begins to move. The data collected from different individuals exhibit significant differences in this field, as different individuals can exhibit differences in behavioral habits. Even when executing the same type of activity, the distribution of data collected from different individuals can vary.

The task of human behavior recognition often involves the use of one or more sensors to collect data, from which relevant features are extracted and inputted into the classification model along with motion category labels to complete supervised learning. However, the model obtained through this process only applies to a scenario where the training and test data follow the same distribution. Thus, due to the limitations imposed by individual differences, the model's limitation is that it can only be used by a specific user. If directly applied to a scenario involving multiple individuals, the individual differences in sensor signals will lead to a significant decrease in the accuracy of movement intention perception. When the movement intention perception task involves multiple users, it is necessary to complete the entire process for each user, including data collection, data processing, and training the model. This not only increases the cost but also consumes more time and energy from the user, reducing the practicality. Therefore, enhancing the generalization ability of the HAR model and improving its recognition accuracy for scenarios involving multiple individuals are vital for extending the applicability of exoskeleton robots and improving the experience of users who use these devices.

To mitigate the cross-subject issue, we develop a dual-channel HAR model

based on CNN. This model achieves end-to-end human behavior recognition and serves as a foundation for subsequent cross-subject transfer. Subsequently, a cross-subject HAR algorithm is proposed. By introducing concepts such as confidence threshold and sparse inquiry, this model gains the ability to recognize uncertain samples, query real labels, and propagate to neighboring samples, thereby constructing a labeled dataset belonging to the current new individual. After training the model by this dataset, the algorithm can adapt to the current new individual, thereby improving behavior recognition accuracy. The main contributions of this paper are summarized as follows:

1) This paper proposes a novel cross-subject HAR method. This method utilizes an active inquiry mechanism to construct a small amount of labeled data belonging to new individuals for model training. This method significantly enhances the model's ability for cross-subject adaptation and reduces the requirement for labeled data.

2) We design a dual-channel model based on CNN. It consists of a temporal feature extractor and a spatial feature extractor, which enables the model learning domain invariant representation for cross-subject adaptation.

3) The proposed method in this paper is validated on two publicly available datasets in the field. Evaluation metrics such as cross-subject recognition accuracy and the proportion of inquiry samples are compared with other existing methods. Experimental results demonstrate that the method proposed in this paper achieves the best cross-subject recognition performance.

2. Related Work

2.1. Human Activity Recognition

Exoskeleton robots operate normally by identifying the user's movement category and selecting the corresponding control strategy to provide assistance, ensuring that the assistance provided to the user is appropriate and timely. Therefore, in order for the exoskeleton robot to provide reasonable assistance and for the user to have a good human-machine interaction experience, it is necessary to establish a human-machine interface to accurately recognize and predict the user's movement intent.

The field of HAR can be traced back to the early 1990s, as demonstrated by the research of Foerster *et al.* [6], which showed that through the process of controlling data collection, the use of high-quality datasets can achieve an accuracy rate of over 95% in recognizing movement intentions. With the rapid development of artificial intelligence technology and the widespread popularity of wearable devices, the accuracy of HAR has been significantly improved, and its application fields have also been greatly expanded. Atchuluun *et al.* [7] proposed a behavior recognition technology based on fuzzy systems, which combines behavior prediction and recognition using data obtained from dual cameras, visible light, and far-infrared light, enabling the recognition of eleven different human behaviors. Ji *et al.* [8] proposed a motion recognition method insensitive

to duration and applicable to complex backgrounds by embedding skeleton information into depth maps obtained by depth cameras, while Jalal *et al.* [9] proposed an online recognition method based on depth video data and human skeletal joint features. Xu *et al.* [10] provided an effective solution for hand motion detection based on mobile first-person view depth sequences. Oyedotun *et al.* [11] improved the accuracy in complex gesture classification tasks using CNN and stacked denoising autoencoders. Pigou *et al.* [12] explored gesture recognition tasks in videos, proposing an end-to-end neural network architecture that achieved better results. Qi *et al.* [13] proposed a dual-layer recognition framework using acceleration data, capable of classifying aerobic, sedentary, and free weight activities, supporting a wider range of physical activity recognition compared to existing methods. Aviles-Cruz *et al.* [14] applied Granger causality theory to motion recognition and proposed a framework for analyzing and classifying individual user activities using three-axis accelerometers in smartphones. With the advent of an aging society, HAR has provided convenience for various aspects such as home care, postoperative rehabilitation, exercise and fitness, stroke detection, epilepsy and Parkinson's research, and monitoring the physical function of the elderly [15].

2.2. Wearable Sensor-Based Human Activity Recognition

Wearable sensor-based HAR has the advantages of convenience, signal stability, and high safety, with no restrictions on human movement range. With the development of smartwatches, smart wristbands, an increasing variety of motion sensors are being integrated into wearable devices, and the development of the Internet of Things and cloud computing [16] has provided conditions for the development of sensor-based HAR. Wearable sensors include motion physiological signal sensors and motion sensors. Physiological sensors are used to collect physiological signals during human movement, such as electromyography (EMG) and mechanomyography (MMG). In this paper, we focus on motion sensors, such as accelerometers and gyroscopes.

Hegde *et al.* [17] designed a wrist-worn sensor containing accelerometers and gyroscopes, achieving a recognition accuracy of over 94% on a set of daily life activities with this device. Chung *et al.* [18] validated that four sensors on the wrists, right ankle, and waist can achieve a recognition accuracy of 91.2% in daily life activities and verified that only two sensors on the left wrist and right ankle can also achieve sufficiently high accuracy. Laput *et al.* [19] classified gestures such as light tapping, clapping, and tapping by improving the sampling rate of accelerometers in smartwatches to collect bioacoustics signals. Muscle contraction mechanical signals can be collected using microphones or accelerometers and used to detect movement patterns. Pham *et al.* [20] developed electronic shoes with miniature wireless accelerometers embedded in the insoles, achieving an average recognition accuracy of 93% for seven daily activities. Wang *et al.* [21] studied the possibility of identifying human movements

using accelerometers and gyroscopes built into smartphones and proposed a feature selection method with better temporal performance. Zhou *et al.* [22] proposed a semi-supervised deep learning framework that effectively utilizes weakly labeled sensor data based on accelerometer data. They designed a distance-based reward rule to address the problem of insufficient sample labeling and developed an intelligent automatic labeling scheme based on deep Q networks. Li *et al.* [23] proposed a method to identify basic activities and transition activities in continuous sensor data streams, accurately distinguishing between segments of adjacent basic activities to determine whether they are transition activities or interference. Chen *et al.* [24] collected accelerometer data using smartphones and segmented activity units, then characterized them using time domain, frequency domain, and wavelet domain features, achieving an accuracy rate of 95.95% for activity classification. Lawal *et al.* [25] used two sensors located at the waist to collect hip movement signals and converted them to the frequency domain, achieving higher accuracy than another state-of-the-art method in predicting human activities.

2.3. Cross-Subject Human Activity Recognition

There are significant differences among individuals and differences in age, gender, physical condition, and behavioral habits can lead to significant differences in signals collected by sensors among different individuals. Therefore, sensor-based HAR models only have good effects for single users, and there will be a significant decrease in performance for new users. There are two methods for achieving cross-subject HAR: the first is to appropriately increase the amount of labeled data from new individuals and train the model, which belongs to semi-supervised learning; the second is to extract features that can achieve motion pattern recognition and are independent of individuals without any labeled data, which belongs to unsupervised learning.

Kongsil *et al.* [26] developed a cross-subject HAR model S-PAR that fuses data from two sensors on smartwatches. This model can recognize activities based on data collected by smartwatches without the need to collect initial labeled data from model users. Gholamiangonabadi *et al.* [27] evaluated six feed-forward neural networks and CNN with two convolutions and one-dimensional filters, as well as four preprocessing scenarios, using leave-one-subject-out cross-validation. Leite *et al.* [28] increased the diversity of participants by generating additional training data that mimic other human subjects in an adversarial learning manner and made the classifier ignore information related to individuals in motion data. Ye *et al.* [29] analyzed the feature confusion problem in cross-subject HAR and summarized it as decision boundary confusion and overlap confusion, which were addressed by optimizing features and decision boundaries and introducing minimal class confusion loss, respectively. Soleimani *et al.* [30] proposed a cross-subject transfer method SA-GAN based on generative adversarial networks, which outperforms all other comparison

methods in over 66% of experiments and ranks second in performance in the remaining 25% of experiments. Wang *et al.* [31] proposed sub-domain adaptation to achieve cross-domain activity recognition by utilizing knowledge from a source domain auxiliary dataset to build models for the target domain. Zhao *et al.* [32] proposed a method of local domain adaptation, which first classifies various activities into static activity clusters and dynamic activity clusters based on the hierarchical structure of human activities, and then aligns the source domain clusters with the target domain clusters. Kongsil *et al.* [33] proposed a cross-subject recognition framework that first classifies activities into dynamic activities and static activities, and then recognizes models for dynamic and static activities using accelerometer and gyroscope data. Suh *et al.* [34] proposed a subject-independent feature extraction method based on adversarial autoencoder structures and maximum mean discrepancy, which learns embedding feature representations independent of individuals from multiple source domain individual datasets and applies these features to motion recognition of target individuals. Kumar *et al.* [35] developed a knowledge transfer-based model DeepTransHAR to address differences between source and target domains in cross-subject domain, cross-sensor problems, using gated recurrent units as memory units to discover and memorize activity patterns in sensor data streams.

Lin *et al.* [36] proposed a training method that only requires a small amount of additional labeled data to improve the generalization ability of pre-trained models. This method balances the efficiency of the nearest class average classifier and the flexibility of cosine similarity. Cruciani *et al.* [37] proposed a method that first identifies a subset of individuals from all source-domain individuals that are most suitable for the target individual and trains a classifier using the data from this subset. Finally, a small amount of labeled data from the target individual is used to update the classifier to adapt it to the specific target individual. The proposed method achieved an F-score of 74.4%, higher than other methods which achieved 70.9%. Soleimani *et al.* [38] proposed a general semi-supervised method based on an adversarial learning framework, which utilizes labeled data from the source domain and unlabeled data from the target individual to address individual differences. Xu *et al.* [39] introduced a fast and robust hybrid model that utilizes domain-adaptive neural networks and deep domain confusion networks to reduce domain shift caused by individual and position variations, and employs a classifier based on online sequential extreme learning machine to quickly learn from a small amount of labeled data in the target domain to update parameters. Zeng *et al.* [40] proposed two semi-supervised methods based on CNN to learn discriminative hidden features from labeled and unlabeled data, as well as raw sensor data, achieving up to an 18% improvement in F1 score on three datasets. Bettini *et al.* [41] proposed a method called FedHAR, which combines semi-supervised learning and federated learning. It incorporates active learning and label propagation to semiautomatically label sensor data for each user, while also including a transfer strategy to provide personalized models for each user. Liu *et al.* [42] proposed

a manifold regularized dynamic graph convolutional network (MRDGCN), which can automatically update structural information through manifold regularization until the model is fitted. Chen *et al.* [43] designed a cooperative training framework that combines attention mechanisms with recursive convolutional models. It selects trustworthy samples from the independent feature spaces of two classifiers, assigns estimated labels to them, and adds them to the training set. Narasimman *et al.* [44] introduced the Mean Teacher semi-supervised model into the field of human motion intent perception, which averages the weights during training to obtain more robust results.

As unsupervised learning does not utilize any labeled data, it requires more computational resources and time to process and infer a large amount of unlabeled data, resulting in lower efficiency and performance compared to semi-supervised learning. In contrast, semi-supervised learning can mitigate the conflict between high data labeling costs and limited performance of unsupervised learning by fully utilizing sparse labeled data to train models.

3. Methods

This section defines the cross-subject issue and introduces our proposed method in detail. The method includes the construction of a HAR model, the pre-training process, the selection of samples with high confidence level, the definition of the clustering loss and the process of model updating.

3.1. Proposed Method

In this section, we propose a transfer algorithm based on confidence threshold, sparse inquiry and label propagation to improve the accuracy of human motion intent perception in cross-subject scenarios and mitigate the impact of individual differences. This algorithm selects samples with relatively high confidence by introducing a confidence threshold and obtains a small number of completely true sample labels through sparse inquiry. Finally, label propagation is used to maximize the benefits of inquiry. **Figure 1** illustrates the overview of the proposed framework for cross-subject HAR, which consists the following steps.

- 1) Train the motion intent perception model on the source domain dataset and save parameters of the pre-trained model that obtaining the highest performance.
- 2) Re-input all samples in the target domain training set into the above pre-trained model, and save the feature vectors, classification results, and classification confidences of each sample output by the feature extractor.
- 3) For each activity, select samples whose confidence degree is greater than the confidence threshold of 0.3 and calculate the average feature vector of these selected samples. The vector is regarded as the center of this activity in the feature space.
- 4) Calculate the Euclidean distance between samples' feature vector and centers of activities. For each sample, compute the ratio of its distance to the nearest center and second nearest center. The closer this ratio is to one, the closer the sample

is to the boundary between two classes.

5) Select the top ten samples with the maximum of the above ratio for each motion class as boundary samples, query the true labels of these boundary samples from the target domain label library, and propagate these labels to the neighboring samples in the feature space.

6) Merge the propagated sample labels with the pseudo-labels of high-confidence center samples to construct a labeled training set on the target domain.

After constructing the above labeled training set, the pretrained model is updated and tested. In the training phase, the cross-entropy loss function and the Adam optimizer are used to update the network.

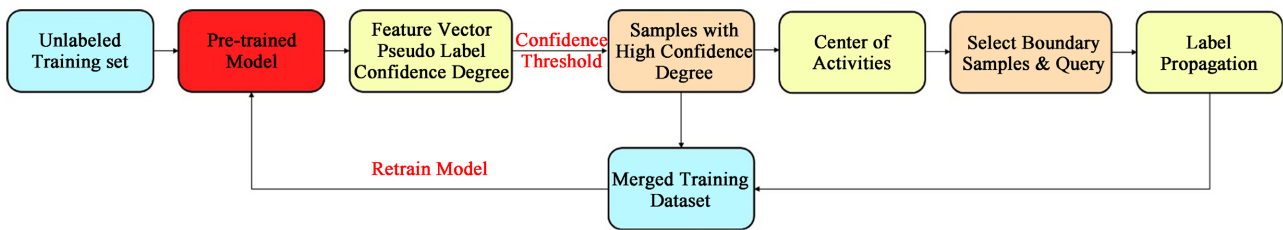


Figure 1. Overview of the proposed framework.

3.2. CNN-Based HAR Model

Since the human body usually moves continuously during a certain activity, the signals collected by IMU sensors have temporal correlation. Moreover, when performing an activity, it usually involves several different parts of the body simultaneously, so the signals collected from different parts of the body also have spatial correlations. So, two-dimensional data containing spatiotemporal information can be constructed as input data with temporal signals as rows and data from different sensors as columns, which can be processed by CNN. The model proposed in this paper is shown in **Figure 2**. As the temporal information mainly exists in the row of the input data, so a rectangular convolution kernel is designed to extract temporal features that characterize the features between rows. The spatial information is mainly contained between the columns, so information from sensors worn on different body parts can be extracted by the spatial feature extractor. The outputs of the spatial and temporal feature extractors are merged and input to the subsequent classifier. The classifiers are composed of fully connected layers, batch normalization layers, and activation function layers.

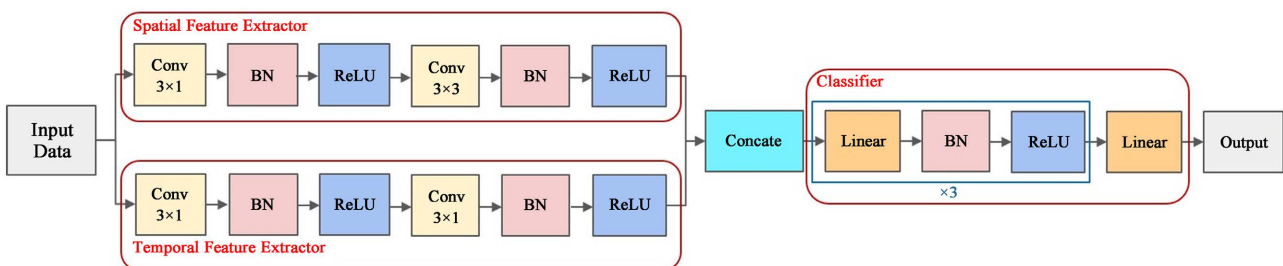


Figure 2. Architecture of the dual-channel CNN model.

4. Experiment

4.1. Datasets

To validate the effectiveness of the CNN model, the following two publicly available datasets are employed in this study, which are widely used in the field of HAR. **Figure 3** illustrates the sensor placement in the DSAD dataset and the PAMAP2 dataset.

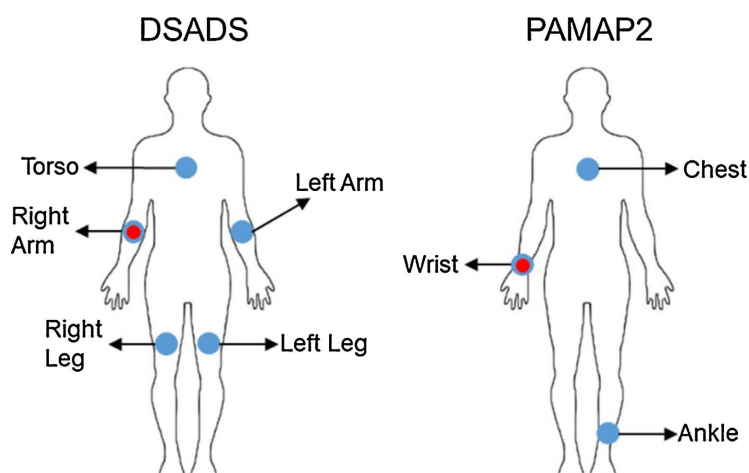


Figure 3. Accuracy and F1-score comparison of activity recognition on DSADS.

4.1.1. Daily and Sports Activities Data Set (DSADS)

This dataset [45] consists of 8 subjects aged between 20 and 30 years. Each subject wears one XsensMTx unit on the trunk, right arm, left arm, right leg, and left leg, respectively. Each unit contains accelerometers, gyroscopes, and magnetometers with a sampling frequency of 25 Hz. There are 19 activities in DSAD (sitting, standing, lying on back, lying on right-side, ascending stairs, descending stairs, standing in an elevator, moving around in an elevator, walking in a parking lot, walking on a treadmill with a speed of 4 km/h in flat, walking on a treadmill with a speed of 4 km/h in 15 deg inclined positions, running at 8 km/h, stepper exercise, cross trainer exercise, cycling in a horizontal position, cycling in a vertical position, rowing, jumping, playing basketball). These subjects perform each activity for 5 minutes.

4.1.2. Physical Activity Monitoring for Aging People (PAMAP2)

This dataset [46] includes 9 subjects, consisting of 8 males and 1 female. Each subject wears 3 IMU sensors on the wrist, chest, and ankle, with a sampling frequency of 100Hz. All subjects perform 12 activities (lying, sitting, standing, walking, running, cycling, Nordic walking, ironing, vacuum cleaning, rope jumping, walking upstairs, walking downstairs).

Preprocessing of the dataset involves data extracting, handling missing values, resampling, performing sliding window sampling, and merging and storing data samples. Due to wireless transmission of data collected by sensors, data loss may occur during transmission. When handling missing values, it is necessary to first detect the positions of missing values and then interpolate the missing positions

with nearby normal data. For the DSAD dataset, the sampling frequency of the sensors is 25 Hz, and the sampling duration for each activity is 5 minutes, resulting in only 7500 samples for each activity. This sample size is significantly insufficient for training neural networks, thus requiring resampling of sensor data to increase the sampling frequency to 100 Hz. No resampling is needed for the PAMAP2 dataset. After resampling, sliding windows are designed for sliding window sampling to obtain more data samples. In this study, the length of the sliding window is set to 300 ms, and the sliding step is set to 30 ms, with overlapping sliding sampling conducted on the one-dimensional time-series data of the sensors. After sliding window sampling, all sensor data is concatenated into a two-dimensional array with each row representing a sample obtained from sliding window sampling and each column representing the data from each sensor.

4.2. Intermethod Evaluation

In this section, the proposed method and the comparative methods are tested on DSAD and PAMAP2 datasets to evaluate their performance on cross-subject adaptation. All subjects and 12 common activities (sitting, standing, lying on back, lying on right-side, ascending stairs, descending stairs, walking in a parking lot, walking on a treadmill with a speed of 4 km/h in 15 deg inclined positions, running, stepper exercise, cross trainer exercise and jumping) in DSAD dataset are selected for this cross-subject experiment. We also choose 7 subjects and 5 common activities (lying, standing, walking, cycling, walking upstairs) from PAMAP2. Leave-one-subject-out validation method is used in this experiment, which means the data of each subject in the dataset is used as the test set in turn and the data of the remaining individuals are used as the training set. For the DSAD and PAMAP2 datasets, the learning rates of our method are 0.00001 and 0.0000001, respectively. The corresponding batch sizes for these datasets are 512 and 256. The comparative methods are briefly introduced as follows.

1) Baseline Method: In practical applications, pre-trained models will be used for entirely new users. Therefore, in the baseline method, the model constructed in Chapter 2 is first trained on the source domain training set and then directly tested on the target domain test set, thus utilizing data that the model has never learned to test the model's generalization ability in the task of cross-subject HAR. When training on the DSAD dataset, the hyperparameters are set as follows: training epochs are 30, batch size is 4096, and learning rate is 0.00001; when training on the PAMAP2 dataset, the hyper-parameters are set as follows: training epochs are 30, batch size is 1024, and learning rate is 0.00001. The test accuracy of the baseline method is regarded as the benchmark, and the test accuracies of other methods are compared with this benchmark, with the improvement in accuracy calculated as the difference.

2) Fine-Tuning Method: Fine-tuning the model after pretraining is a commonly used transfer learning method, which consists of a pre-training stage and a fine-tuning stage. The purpose of pre-training the model is to allow the model to learn

the necessary knowledge for the task from existing data beforehand so that when performing similar tasks again, it can obtain better results more quickly. The fine-tuning stage involves retraining the pre-trained model on a new dataset and updating some parameters to make them more suitable for the new dataset, thus introducing fine-tuning after pre-training can improve the performance of pre-trained models on new tasks. The specific steps of this method are as follows:

a) Construct the network structure and randomly initialize parameters.

b) Pre-training stage: Train the model on the source domain training set dataset, the hyper-parameters are set as follows: training epochs are 30, batch size is 4096, and learning rate is 0.00001; when training on the PAMAP2 dataset, the hyper-parameters are set as follows: training epochs are 30, batch size is 1024, and learning rate is 0.00001.

c) Fine-tuning stage: After obtaining the pre-trained source model, fix the parameters of its feature extraction part and no longer update them. Initialize the classifier part randomly and update it on the target domain training set. When fine-tuning on the DSAD dataset, the hyper parameters are set as follows: training epochs are 50, batch size is 512, and learning rate is 0.00000001. When fine-tuning on the PAMAP2 dataset, the hyper-parameter settings are as follows: the training epochs are 50, the batch size is 512, and the learning rate is 0.00000001. After the model is personalized for the target subject, it is evaluated using the test set of the target subject and the recognition accuracy is calculated.

3) Unsupervised Cross-Subject Adaptation for Predicting Human Locomotion Intent [47]: This method is a classic domain adaptation method within the field of HAR, which is based on Maximum Classifier Discrepancy (MCD). The idea of MCD is to have the feature generator and two classifiers contend with each other during training, thereby changing the distribution of the target domain samples and enabling the classifier to classify the target domain samples well. The training epochs for both datasets are 30. The learning rate for DSAD and PAMAP2 are 0.00001 and 0.00000001, respectively. The batch sizes for DSAD and PAMAP2 are 512 and 128, respectively.

4.3. Experimental Results

On the DSAD dataset, with 8 subjects taking turns as the target domain individuals, and the data of the remaining people as the source domain, the algorithm iterates for five rounds. The test results of the proposed method after each round of iteration are shown in **Table 1**. According to the results in the table, the following conclusions can be drawn: the method proposed in this paper can achieve the best performance on all target individuals, and the test accuracy gradually increases with the number of iterations.

The comparison of the test accuracy improvement of the method proposed in this paper with other methods is shown in **Table 2**. For the 8 target individuals, the change in test accuracy of the method proposed in this paper compared to the baseline method is 2.30%, 7.89%, 6.00%, 3.98%, 3.86%, 1.05%, 2.44%, 10.54%,

respectively. Compared to the finetuning method, the change in test accuracy is 0.19%, -0.72%, 1.71%, 0.37%, -0.24%, 1.64%, 0.00%, 0.17%, respectively. Compared to the MCD method, the change in test accuracy is 0.27%, 4.05%, 1.11%, 16.56%, -1.17%, -2.39%, 3.09%, 0.69%, respectively.

Table 1. Test accuracy of the proposed method on DSAD.

Target Subject	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5
P1	87.50%	87.91%	88.16%	88.34%	88.42%
P2	87.28%	87.48%	87.72%	87.85%	88.09%
P3	82.16%	83.14%	83.76%	84.50%	84.95%
P4	92.49%	93.58%	94.40%	94.64%	94.98%
P5	94.45%	94.75%	94.93%	95.02%	95.06%
P6	93.72%	93.29%	93.15%	93.16%	93.47%
P7	90.58%	90.94%	91.25%	91.54%	91.54%
P8	96.82%	97.06%	97.03%	97.12%	97.20%

Table 2. Intermethod comparison on DSAD.

Target Subject	Baseline	Our	Fine-tuning	Classifier 1	Classifier 2
P1	86.04%	88.42%	88.23%	88.15%	88.09%
P2	80.20%	88.09%	88.81%	83.06%	84.04%
P3	78.95%	84.95%	83.24%	83.84%	82.91%
P4	91.00%	94.98%	94.61%	78.42%	77.19%
P5	91.20%	95.06%	95.30%	96.23%	95.86%
P6	92.42%	93.47%	91.83%	95.07%	95.36%
P7	89.54%	91.98%	91.98%	88.23%	88.89%
P8	86.66%	97.20%	97.03%	96.51%	96.23%

On the PAMAP2 dataset, with 6 subjects taking turns as the target domain individuals, and the data of the remaining people as the source domain, the algorithm iterates for 5 rounds. The test results of the method proposed in this paper are shown in **Table 3**. The method proposed in this paper achieves a test accuracy greater than 70% on all target individuals. With the increase in the number of iterations, the test accuracy shows an upward trend on most target individuals. The comparison of the test accuracy improvement of the method proposed in this paper with other methods is shown in **Table 4**. For these target subjects, the changes in test accuracy of our method compared to the baseline method are 22.70%, 33.10%, 41.90%, 21.35%, 20.41%, 60.30%, respectively. Compared to the fine-tuning method, the changes in test accuracy are -0.68%, 0.06%, 0.40%, 1.64%, 0.87% and 0.43%, respectively; compared to the MCD method, the changes in test accuracy are -3.02%, 2.59%, 12.59%, 6.78%, 8.14%, and 2.52%, respectively.

Table 3. Test accuracy of the proposed method on PAMAP2.

Target Subject	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5
P1	70.72%	70.10%	70.83%	69.91%	70.50%
P2	84.84%	85.22%	85.54%	85.57%	86.15%
P3	85.99%	85.94%	86.55%	86.38%	86.98%
P4	73.78%	73.92%	77.44%	79.29%	79.35%
P5	77.37%	78.99%	80.54%	80.84%	81.49%
P6	80.26%	80.28%	81.01%	80.98%	81.34%

Table 4. Intermethod comparison on PAMAP2.

Target Subject	Baseline	Our	Fine-tuning	Classifier 1	Classifier 2
P1	48.02%	70.50%	71.18%	62.98%	73.52%
P2	53.05%	86.15%	86.10%	83.57%	83.22%
P3	45.08%	86.98%	86.58%	63.42%W	74.39%
P4	58.00%	79.35%	77.71%	62.37%	72.75%
P5	61.08%	81.49%	80.62%	73.35%	46.79%
P6	21.04%	81.34%	80.91%	78.82%	57.95%

5. Discussion

In this section, we further analyze the results of the aforementioned cross-subject movement intention perception experiments. To measure the additional burden on users caused by active querying of unlabeled samples in our method, we define the query sample ratio as the ratio of the total number of queried samples to the total number of target domain samples. For the DSAD dataset, compared to the baseline method, finetuning method, and MCD method, the average changes in test accuracy on all target subjects are 4.76%, 0.57%, and 2.78%, respectively. The results show that compared to the baseline and MCD methods, our method has improved accuracy in cross-subject movement intention perception tasks and is very close to the fine-tuning method, which is the upper limit. In our method, the average query sample ratio for all target subjects is 1.21%, while the fine-tuning method uses 42.86% of the labeled data in the target domain. On the PAMAP2 dataset, our method's average test accuracy exceeds the baseline method, fine-tuning method, and MCD method by 33.29%, 0.68%, and 4.93%, respectively. Therefore, our method further improves the accuracy of cross-subject recognition compared to the baseline and MCD methods, and is close to the effect of the fine-tuning method. The average query sample ratio for our method on six target subjects is only 1.08%, far less than the number of labeled data required in the target domain by the fine-tuning method. The proposed method outperforms MCD mainly because our method is refined using labeled target domain samples, which enables the method to learn the data distribution of target domain in a supervised

manner. Moreover, the label propagation extends the number of labeled samples in the target domain without bring heavy burden to users like fine-tuning. According to the above results, our method can not only improve the accuracy of cross-subject HAR but also reduce the burden on users.

6. Conclusion

To mitigate the cross-subject issue and enable lower limb exoskeleton robots adapting to new users, this paper proposes a cross-subject HAR method that consists of confidence threshold, sparse querying, and label propagation. This method is compared with the baseline method, MCD method, and fine-tuning method on two publicly datasets in the field. The experimental results show that the proposed method achieves the highest recognition accuracy on cross-subject adaptation task and is close to the fine-tuning method, which is seen as the upper bound. Moreover, the number of samples required for active querying in our method is far less than that in the fine-tuning method, significantly reducing the burden on users in the querying and annotating process, which improves the user experience.

7. Future Work

This work focuses on the cross-subject scenario in HAR domain and mitigates individual difference by proposing our method. An interesting extension of this work is to take unknown activities performed by target subjects into consideration. In that case, individual difference and class difference exists between source domain and target domain simultaneously. The HAR model would be able to adapt to this challenging case by enhancing its generalization ability.

Funding

This work is partly funded by Technique Program of Jiangsu (No. BE2021086).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Hayashi, T., Kawamoto, H. and Sankai, Y. (2005) Control Method of Robot Suit HAL Working as Operator's Muscle Using Biological and Dynamical Information. 2005 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, 2-6 August 2005, 3063-3068. <https://doi.org/10.1109/iros.2005.1545505>
- [2] Dollar, A.M. and Herr, H. (2008) Design of a Quasi-Passive Knee Exoskeleton to Assist Running. 2008 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, 22-26 September 2008, 747-754. <https://doi.org/10.1109/iros.2008.4651202>
- [3] Wang, S., Wang, L., Meijneke, C., van Asseldonk, E., Hoellinger, T., Cheron, G., *et al.* (2015) Design and Control of the MINDWALKER Exoskeleton. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **23**, 277-286. <https://doi.org/10.1109/tnsre.2014.2365697>

- [4] Krut, S., Benoit, M., Dombre, E. and Pierrot, F. (2010). Moonwalker, a Lower Limb Exoskeleton Able to Sustain Bodyweight Using a Passive Force Balancer. 2010 *IEEE International Conference on Robotics and Automation*, Anchorage, 3-7 May 2010, 2215-2220. <https://doi.org/10.1109/robot.2010.5509961>
- [5] Ergasheva, B.I. (2017) Lower Limb Exoskeletons: Brief Review. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, **17**, 1153-1158. <https://doi.org/10.17586/2226-1494-2017-17-6-1153-1158>
- [6] Foerster, F. and Smeja, M. (1999) Joint Amplitude and Frequency Analysis of Tremor Activity. *Electromyography and Clinical Neurophysiology*, **39**, 11-19.
- [7] Batchuluun, G., Kim, J.H., Hong, H.G., Kang, J.K. and Park, K.R. (2017) Fuzzy System Based Human Behavior Recognition by Combining Behavior Prediction and Recognition. *Expert Systems with Applications*, **81**, 108-133. <https://doi.org/10.1016/j.eswa.2017.03.052>
- [8] Ji, X., Cheng, J., Feng, W. and Tao, D. (2018) Skeleton Embedded Motion Body Partition for Human Action Recognition Using Depth Sequences. *Signal Processing*, **143**, 56-68. <https://doi.org/10.1016/j.sigpro.2017.08.016>
- [9] Jalal, A., Kim, Y., Kim, Y., Kamal, S. and Kim, D. (2017) Robust Human Activity Recognition from Depth Video Using Spatiotemporal Multi-Fused Features. *Pattern Recognition*, **61**, 295-308. <https://doi.org/10.1016/j.patcog.2016.08.003>
- [10] Xu, C., Govindarajan, L.N. and Cheng, L. (2017) Hand Action Detection from Ego-Centric Depth Sequences with Error-Correcting Hough Transform. *Pattern Recognition*, **72**, 494-503. <https://doi.org/10.1016/j.patcog.2017.08.009>
- [11] Oyedotun, O.K. and Khashman, A. (2016) Deep Learning in Vision-Based Static Hand Gesture Recognition. *Neural Computing and Applications*, **28**, 3941-3951. <https://doi.org/10.1007/s00521-016-2294-8>
- [12] Pigou, L., van den Oord, A., Dieleman, S., Van Herreweghe, M. and Dambre, J. (2016) Beyond Temporal Pooling: Recurrence and Temporal Convolutions for Gesture Recognition in Video. *International Journal of Computer Vision*, **126**, 430-439. <https://doi.org/10.1007/s11263-016-0957-7>
- [13] Qi, J., Yang, P., Hanneghan, M., Tang, S. and Zhou, B. (2019) A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors. *IEEE Internet of Things Journal*, **6**, 1384-1393. <https://doi.org/10.1109/jiot.2018.2846359>
- [14] Aviles-Cruz, C., Rodriguez-Martinez, E., Villegas-Cortez, J. and Ferreyra-Ramirez, A. (2019) Granger-Causality: An Efficient Single User Movement Recognition Using a Smartphone Accelerometer Sensor. *Pattern Recognition Letters*, **125**, 576-583. <https://doi.org/10.1016/j.patrec.2019.06.029>
- [15] Wang, Y., Cang, S. and Yu, H. (2019) A Survey on Wearable Sensor Modality Centred Human Activity Recognition in Health Care. *Expert Systems with Applications*, **137**, 167-190. <https://doi.org/10.1016/j.eswa.2019.04.057>
- [16] Dang, L.M., Piran, M.J., Han, D., Min, K. and Moon, H. (2019) A Survey on Internet of Things and Cloud Computing for Healthcare. *Electronics*, **8**, Article 768. <https://doi.org/10.3390/electronics8070768>
- [17] Hegde, N., Bries, M., Swibas, T., Melanson, E. and Sazonov, E. (2018) Automatic Recognition of Activities of Daily Living Utilizing Insole-Based and Wrist-Worn Wearable Sensors. *IEEE Journal of Biomedical and Health Informatics*, **22**, 979-988. <https://doi.org/10.1109/jbhi.2017.2734803>
- [18] Chung, S., Lim, J., Noh, K.J., Gue Kim, G. and Jeong, H.T. (2018) Sensor Positioning

- and Data Acquisition for Activity Recognition Using Deep Learning. 2018 *International Conference on Information and Communication Technology Convergence*, Jeju, 17-19 October 2018, 154-159. <https://doi.org/10.1109/ictc.2018.8539473>
- [19] Laput, G., Xiao, R. and Harrison, C. (2016) ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, Toko, 16-19 October, 321-333. <https://doi.org/10.1145/2984511.2984582>
- [20] Pham, C., Diep, N.N. and Phuong, T.M. (2017) E-Shoes: Smart Shoes for Unobtrusive Human Activity Recognition. 2017 *9th International Conference on Knowledge and Systems Engineering (KSE)*, Hue, 19-21 October 2017, 269-274. <https://doi.org/10.1109/kse.2017.8119470>
- [21] Wang, A., Chen, G., Yang, J., Zhao, S. and Chang, C. (2016) A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone. *IEEE Sensors Journal*, **16**, 4566-4578. <https://doi.org/10.1109/jsen.2016.2545708>
- [22] Zhou, X., Liang, W., Wang, K.I., Wang, H., Yang, L.T. and Jin, Q. (2020) Deep-Learning-Enhanced Human Activity Recognition for Internet of Healthcare Things. *IEEE Internet of Things Journal*, **7**, 6429-6438. <https://doi.org/10.1109/jiot.2020.2985082>
- [23] Li, J., Tian, L., Wang, H., An, Y., Wang, K. and Yu, L. (2019) Segmentation and Recognition of Basic and Transitional Activities for Continuous Physical Human Activity. *IEEE Access*, **7**, 42565-42576. <https://doi.org/10.1109/access.2019.2905575>
- [24] Chen, Y. and Shen, C. (2017) Performance Analysis of Smartphone-Sensor Behavior for Human Activity Recognition. *IEEE Access*, **5**, 3095-3110. <https://doi.org/10.1109/access.2017.2676168>
- [25] Lawal, I.A. and Bano, S. (2019) Deep Human Activity Recognition Using Wearable Sensors. *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, Rhodes, 5-7 June 2019, 45-48. <https://doi.org/10.1145/3316782.3321538>
- [26] Kongsil, K., Suksawatchon, J. and Suksawatchon, U. (2019) Physical Activity Recognition Using Streaming Data from Wrist-Worn Sensors. 2019 *4th International Conference on Information Technology*, Bangkok, 24-25 October 2019, 274-279. <https://doi.org/10.1109/incit.2019.8912130>
- [27] Gholamiangonabadi, D., Kiselov, N. and Grolinger, K. (2020) Deep Neural Networks for Human Activity Recognition with Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection. *IEEE Access*, **8**, 133982-133994. <https://doi.org/10.1109/access.2020.3010715>
- [28] Leite, C.F.S. and Xiao, Y. (2020) Improving Cross-Subject Activity Recognition via Adversarial Learning. *IEEE Access*, **8**, 90542-90554. <https://doi.org/10.1109/access.2020.2993818>
- [29] Ye, Y., Zhou, Q., Pan, T., Huang, Z. and Wan, Z. (2021) Alleviating Feature Confusion in Cross-Subject Human Activity Recognition via Adversarial Domain Adaptation Strategy. 2021 *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, Mexico, 1-5 November 2021, 7586-7589. <https://doi.org/10.1109/embc46164.2021.9630655>
- [30] Soleimani, E. and Nazerfard, E. (2021) Cross-Subject Transfer Learning in Human Activity Recognition Systems Using Generative Adversarial Networks. *Neurocomputing*, **426**, 26-34. <https://doi.org/10.1016/j.neucom.2020.10.056>
- [31] Lu, W., Chen, Y., Wang, J. and Qin, X. (2021) Cross-Domain Activity Recognition via Substructural Optimal Transport. *Neurocomputing*, **454**, 65-75. <https://doi.org/10.1016/j.neucom.2021.04.124>

- [32] Zhao, J., Deng, F., He, H. and Chen, J. (2021) Local Domain Adaptation for Cross-Domain Activity Recognition. *IEEE Transactions on Human-Machine Systems*, **51**, 12-21. <https://doi.org/10.1109/thms.2020.3039196>
- [33] Kongsil, K., Suksawatchon, J. and Suksawatchon, U. (2020) Wrist-Worn Physical Activity Recognition: A Fusion Learning Approach. 2020 *5th International Conference on Information Technology*, Chonburi, 21-22 October 2020, 116-121. <https://doi.org/10.1109/incit50588.2020.9310980>
- [34] Suh, S., Rey, V. and Lukowicz, P. (2021) Adversarial Deep Feature Extraction Network for User Independent Human Activity Recognition.
- [35] Kumar, P. and Suresh, S. (2023) Deeptranshar: A Novel Clustering-Based Transfer Learning Approach for Recognizing the Cross-Domain Human Activities Using Grus (Gated Recurrent Units) Networks. *Internet of Things*, **21**, Article 100681. <https://doi.org/10.1016/j.iot.2023.100681>
- [36] Lin, C. and Marculescu, R. (2020) Model Personalization for Human Activity Recognition. 2020 *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Austin, 23-27 March 2020, 1-7. <https://doi.org/10.1109/percomworkshops48775.2020.9156229>
- [37] Cruciani, F., Nugent, C.D., Quero, J.M., Cleland, I., Mccullagh, P., Synnes, K., *et al.* (2020) Personalizing Activity Recognition with a Clustering Based Semi-Population Approach. *IEEE Access*, **8**, 207794-207804. <https://doi.org/10.1109/access.2020.3038084>
- [38] Soleimani, E., Khodabandelou, G., Chibani, A. and Amirat, Y. (2022) Generic Semi-Supervised Adversarial Subject Translation for Sensor-Based Activity Recognition. *Neurocomputing*, **500**, 649-661. <https://doi.org/10.1016/j.neucom.2022.05.075>
- [39] Xu, Q., Wei, X., Bai, R., Li, S. and Meng, Z. (2023) Integration of Deep Adaptation Transfer Learning and Online Sequential Extreme Learning Machine for Cross-Person and Cross-Position Activity Recognition. *Expert Systems with Applications*, **212**, Article 118807. <https://doi.org/10.1016/j.eswa.2022.118807>
- [40] Zeng, M., Yu, T., Wang, X., Nguyen, L.T., Mengshoel, O.J. and Lane, I. (2017) Semi-supervised Convolutional Neural Networks for Human Activity Recognition. 2017 *IEEE International Conference on Big Data*, Boston, 11-14 December 2017, 522-529. <https://doi.org/10.1109/bigdata.2017.8257967>
- [41] Bettini, C., Civitaresse, G. and Presotto, R. (2021) Personalized Semi-Supervised Federated Learning for Human Activity Recognition.
- [42] Liu, W., Fu, S., Zhou, Y., Zha, Z. and Nie, L. (2021) Human Activity Recognition by Manifold Regularization Based Dynamic Graph Convolutional Networks. *Neurocomputing*, **444**, 217-225. <https://doi.org/10.1016/j.neucom.2019.12.150>
- [43] Chen, K., Yao, L., Zhang, D., Wang, X., Chang, X. and Nie, F. (2020) A Semi-Supervised Recurrent Convolutional Attention Model for Human Activity Recognition. *IEEE Transactions on Neural Networks and Learning Systems*, **31**, 1747-1756. <https://doi.org/10.1109/tnnls.2019.2927224>
- [44] Narasimman, G., Lu, K., Raja, A., Foo, C., Aly, M., Jiang, L. and Chandrasekhar, V. (2021) A*har: A New Benchmark towards Semi-Supervised Learning for Class-Imbalanced Human Activity Recognition.
- [45] Altun, K., Barshan, B. and Tunçel, O. (2010) Comparative Study on Classifying Human Activities with Miniature Inertial and Magnetic Sensors. *Pattern Recognition*, **43**, 3605-3620. <https://doi.org/10.1016/j.patcog.2010.04.019>
- [46] Reiss, A. and Stricker, D. (2012) Introducing a New Benchmarked Dataset for Activity

Monitoring. 2012 *16th International Symposium on Wearable Computers*, Newcastle, 18-22 June 2012, 108-109. <https://doi.org/10.1109/iswc.2012.13>

- [47] Zhang, K., Wang, J., de Silva, C.W. and Fu, C. (2020) Unsupervised Cross-Subject Adaptation for Predicting Human Locomotion Intent. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **28**, 646-657. <https://doi.org/10.1109/tnsre.2020.2966749>