Journal of Computer Science and Frontier Technologies
ISSN Print: 3104-4204
ISSN Online: 3104-4212
CD ISSN: 3104-4220
https://doi.org/10.63313/JCSFT.9104

Generalized Few-Shot Human Activity Recognition Based on Knowledge Transfer and Fine-Tuning Strategies

Bingze Guo
College of Computer Science and Technology, Qingdao University, Qingdao 266071, China
Email: 2660279812@qq.com

How to cite this paper: Guo, B. (2026). Generalized few-shot human activity recognition based on knowledge transfer and fine-tuning strategies. Journal of Computer Science and Frontier Technologies, 4(3), 14–25. ISSN Print: 3104-4204, ISSN Online: 3104-4212.
https://doi.org/10.63313/JCSFT.9104
Published: 2026-10-20

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Abstract

To address the limited recognition capability of novel classes under few-shot conditions and the degradation of existing-class performance after learning novel classes in sensor-based human activity recognition, this paper proposes a knowledge transfer method to improve the overall recognition performance of existing and novel classes. First, a deep learning model is pretrained using abundant training samples from existing classes to learn feature representations and discriminative knowledge. Then, few-shot training samples of novel classes are fed into the pretrained model to construct a similarity matrix between the classifier weights of novel and existing classes. This matrix is used to generate novel-class classifier weights by weighted combination of existing-class classifier weights. Finally, the classifier weights of existing and novel classes are concatenated and jointly fine-tuned to improve recognition across all classes. Experiments are conducted on three public human activity recognition datasets under 1-shot, 5-shot, and 10-shot settings, with comparisons against five existing methods. The harmonic mean of the average classification accuracies of existing and novel classes is adopted as the evaluation metric. Experimental results demonstrate that the proposed method achieves superior overall recognition performance across different datasets and few-shot settings. The results indicate that knowledge transfer based on classifier-weight similarity can effectively exploit the discriminative knowledge of existing classes, improve novel-class recognition, and alleviate the performance degradation of existing classes, thereby achieving a better balance between existing- and novel-class recognition.

Keywords

Deep Learning; Generalized Few-Shot Learning; Human Activity Recognition; Sensor Data

1. Introduction

With the rapid development of wearable devices and the Internet of Things (IoT), sensor-based human activity recognition (HAR) [1-3] has become an important

research direction in the field of artificial intelligence. Traditional HAR methods typically rely on manually designed features and large amounts of labeled data, which involve complex feature extraction and high data annotation costs. In recent years, deep learning has demonstrated promising performance in HAR by automatically extracting features from raw sensor data. However, the high cost of annotating data for novel activities means that only a limited number of labeled samples are typically available for new activity classes in practical applications, making it difficult to meet the large-scale training data requirements of deep learning models. Therefore, few-shot learning has gradually become a research focus in this field.

Existing few-shot learning [4] methods mainly focus on the recognition performance of novel classes. However, in practical applications, models must not only recognize novel classes with limited samples but also maintain their recognition capability for base classes (i.e., existing classes). Consequently, generalized few-shot learning, which aims to simultaneously maintain the recognition performance of both base and novel classes, has attracted increasing attention. Compared with conventional few-shot learning, generalized few-shot learning faces a more pronounced class imbalance problem: base classes typically have a large number of training samples, whereas novel classes have only a few samples. The abundance of base-class data can easily cause the model to develop a class bias, resulting in predictions that tend to favor base classes. Meanwhile, the limited number of novel-class samples restricts the model's ability to sufficiently learn their class-specific characteristics. Therefore, how to exploit knowledge from base classes to improve novel-class recognition while maintaining the recognition performance of base classes is an important problem in generalized few-shot human activity recognition.

To address the above challenges, this paper investigates effective knowledge transfer from base classes to novel classes for generalized few-shot human activity recognition based on sensor data. The objective is to fully exploit the prior knowledge of base classes, mitigate the effects of class imbalance and the limited availability of novel-class samples, and thereby improve the overall recognition performance of the model on both base and novel classes. This study provides a new perspective for human activity recognition under few-shot conditions.

The main contributions of this study are summarized as follows:

- We introduce a Transformer encoder into the feature extraction module and leverage the attention mechanism to effectively capture temporal dependencies, enabling the model to extract more representative feature embeddings from sensor data.
- We propose a knowledge transfer strategy to generate novel-class classification weights based on the classification weights of base classes. In addition, balanced sampling of base-class and novel-class samples is employed during fine-tuning to effectively alleviate class imbalance and the limited availability of novel-class

training data.

- Extensive experiments are conducted on three public datasets and compared with five existing methods. The results demonstrate that the proposed method consistently outperforms existing approaches across different few-shot settings.

2. Related Work

2.1. Few-Shot Activity Recognition

Few-shot learning aims to address the challenge of enabling models to effectively recognize novel classes when only a limited number of training samples are available. Representative approaches include prototypical networks, matching networks, and meta-learning-based methods. Gong et al. [5] proposed an improved convolutional neural network as the feature extraction module and employed a prototypical network for activity recognition to address few-shot human activity recognition based on radar signals. Ding et al. [6] introduced an attention mechanism to address the problem of recognizing few-shot human activities under different initial states using WiFi signals. Shi et al. [7] focused on few-shot activity classification in new environments and proposed a matching network based on enhanced Channel State Information (CSI). Hu et al. [8] proposed a lightweight few-shot learning framework consisting of a feature extraction sub-network and a similarity discrimination sub-network for recognizing novel gesture categories. Reference [9] mainly investigated the application of generative adversarial networks to few-shot activity recognition.

2.2. Generalized Few-Shot Activity Recognition

Compared with conventional few-shot learning (FSL), generalized few-shot learning (GFSL) is more consistent with practical application requirements, as it requires models not only to effectively recognize novel classes but also to maintain satisfactory performance on base classes. Representative studies have mainly focused on improving novel-class recognition, enhancing generalization capability, alleviating data imbalance, and distinguishing similar human activities. Kasnesis et al. [10] proposed an improved relation network that integrates metric learning and a multi-head attention mechanism to address the challenge of recognizing novel classes with only a few available samples. Khan et al. [11] established relationships between activity classes in the source and target domains to reduce the dependence on labeled samples and improve the recognition capability for novel classes in the target domain. Liu et al. [12] designed a weight generator for generalized human activity recognition. Wijayasingha et al. [13] investigated the limited generalization capability of novel classes in generalized few-shot activity recognition. Nguyen et al. [14] introduced semantic knowledge to alleviate the data imbalance problem in generalized activity recognition. Ghorpade et al. [15] focused on the difficulty of distinguishing specific human activities in the generalized few-shot recognition

setting.

2.3. Transformer

In deep learning-based human activity recognition, the feature extraction capability of the model directly affects its recognition performance. Traditional convolutional neural networks (CNNs) have demonstrated strong capabilities in extracting local features; however, their limited receptive fields make it difficult to effectively capture global dependencies in long temporal sequences. Recurrent neural networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM), can capture temporal dependencies in sequential data, but they often suffer from gradient vanishing and low training efficiency when processing long sequences.

In recent years, the Transformer architecture has been widely applied to sequential data analysis due to its powerful global modeling capability. Unlike CNNs, which primarily focus on local feature extraction, Transformers employ a self-attention mechanism, to capture long-range dependencies, making them particularly suitable for sensor data with complex temporal patterns. Therefore, Transformer-based architectures have become an effective choice for extracting discriminative temporal features in HAR. In this paper, a Transformer encoder is introduced into the feature extraction stage to enhance the model's ability to capture critical temporal information. Moreover, its global modeling capability can alleviate feature instability caused by the limited samples of novel classes, thereby improving the overall discriminative capability of the model for both base and novel classes.

3. Proposed Activity Recognition Framework

The proposed method framework consists of three stages: 1) training the feature extractor and the base-class classifier, 2) generating classification weights for novel classes, 3) fine-tuning the joint classifier. The overall training procedure is illustrated in Figure 1. The details of each stage are described as follows.

3.1. Problem Formulation

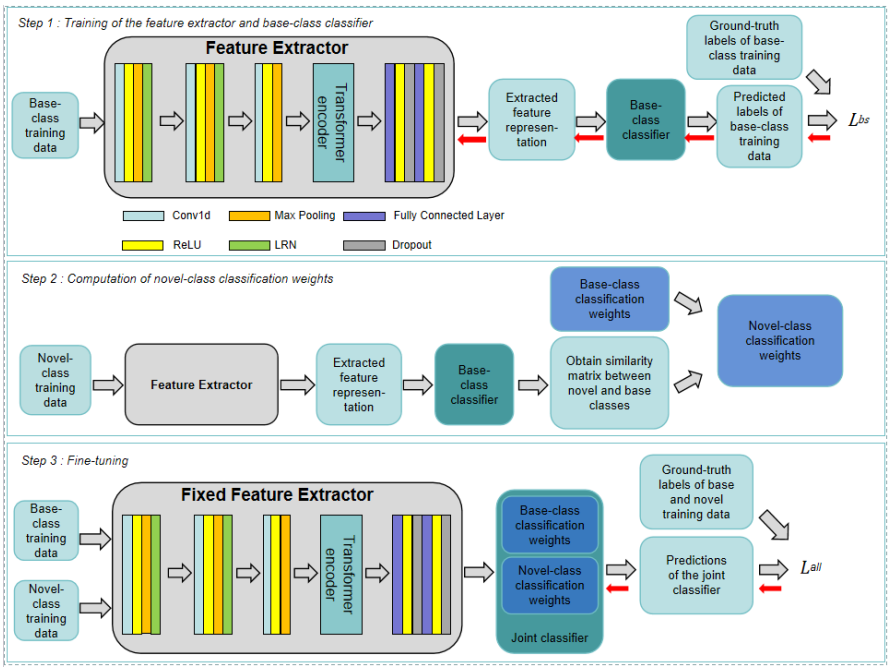
The problem is defined as follows. Given a human activity dataset based on sensor data, denoted as D_t, after a series of preprocessing steps, its label set can be represented as $C = C_b \cup C_n$, where $C_b \cap C_n = \emptyset$. Here, C_b denotes the set of base classes, and C_n denotes the set of novel classes, which are mutually exclusive.

Assume that the base classes contain N categories and the novel classes contain M categories, i.e., $|C_b| = N$ and $|C_n| = M$, Accordingly, the dataset can be divided into a base dataset and a novel dataset: $D_t = D_b \cup D_n$.The base dataset is defined as:$D_b = \{(x_i, y_i) \mid y_i \in C_b\}$, and the novel dataset is defined as: $D_n = \{(x_i, y_i) \mid y_i \in C_n\}$. Under the few-shot setting, each novel class c contains K training samples, i.e., $|\{x_i \mid y_i = c\}| = K, \ \forall c \in C_n$.

The objective of this work is to design a classification model that can accurately

recognize both base and novel classes during the testing phase, i.e., to achieve effective discrimination over the unified label space C.

Figure 1. The overall training procedure of the proposed framework.



3.2. Training of the Feature Extractor and Base-Class Classifier

This stage consists of two components: a convolutional feature extraction module and a fully connected classification module. A convolutional neural network is adopted as the feature extraction module because of its superior performance in HAR tasks.During this stage, the base-class training dataset D_b^{tr} is used as the input to the model.

For the N base classes, an equal number of training samples are randomly selected from each class to alleviate the influence of class imbalance during training. Specifically, for each base class $c \in C_b$, n samples are selected to construct the training subset. Therefore, the total number of training samples used for training is $|D_b^{tr}| = N \times n$.

Given an input sample $x_i \in \mathbb{R}^{T\times C}$, where T denotes the temporal window length and C represents the number of sensor channels, the input is first transposed to $\mathbb{R}^{C\times T}$. Subsequently, three convolutional modules are employed to extract local temporal features: $h_1 = \text{ConvBlock}_1(x_i)$, $h_2 = \text{ConvBlock}_2(h_1)$, and $h_3 = \text{ConvBlock}_3(h_2)$. The first two convolutional modules each consist of a one-dimensional convolutional layer, an activation layer, a max-pooling layer, and a local response normalization layer. The third convolutional module contains a one-dimensional convolutional layer, an activation layer, and a max-pooling layer.One-dimensional convolutions are performed along the temporal dimension

across all sensor channels.After the three convolutional modules, the extracted feature representation is denoted as $h_3 \in \mathbb{R}^{D\times T'}$, where D and T′ represent the feature dimension and temporal length after convolution, respectively.

To capture long-range temporal dependencies, the output of the convolutional modules is transposed into $h_3 \in \mathbb{R}^{T'\times D}$ and then fed into the Transformer encoder to obtain $h' = \text{TransformerEncoder}(h_3)$. The Transformer encoder adopts a multi-head self-attention mechanism, whose core computation is defined as:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{1}$$

where $Q = h_3 W_Q$, $K = h_3 W_K$, and $V = h_3 W_V$. Here, W_Q, W_K and W_V are learnable projection matrices. For an input vector $z = [z_1, z_2, \dots, z_n]$, the Softmax function is defined as:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}, \quad i = 1,2, \dots, n \tag{2}$$

This structure enables the model to capture dependencies between arbitrary positions in the temporal sequence, thereby enhancing the global representation capability of the extracted features. After feature extraction by the convolutional and Transformer encoder modules, two fully connected layers are used to further transform the features and obtain the final feature representation f. The extracted feature f is then fed into a fully connected classification layer to produce the final prediction.

The cross-entropy loss function is adopted as the optimization objective, defined as:

$$\mathcal{L}bs = -\frac{1}{N \times n}\sum_{i=1}^{N\times n}\sum_{j=1}^{N}(\mathbb{I}(y_i = c_j)\ \log(\hat{y}_{i,j})) \tag{3}$$

where $\hat{y}_{i,j}$ denotes the predicted probability that the i-th sample belongs to class c_j, which is obtained from the output of the classification layer and computed via the Softmax function. y_i represents the ground-truth label of the i-th sample, and $\mathbb{I}(\cdot)$ is an indicator function that equals 1 when $y_i = c_j$ and 0 otherwise.After training, the feature extraction module and the base-class classification module are obtained.

3.3. Computation of Novel-Class Classification Weights

In the previous stage, the classification weights of the base classes are obtained as:

$$W_b \in \mathbb{R}^{N\times d} = \{W_1^b, W_2^b, \dots, W_N^b\} \tag{4}$$

where d denotes the feature dimension, and W_i^b represents the classification weight vector of the i-th base class. Therefore, the problem of obtaining classification weights for the novel classes can be formulated as constructing the classification weight matrix of the novel classes $W_n \in \mathbb{R}^{M\times d}$ by utilizing the existing model parameters and a limited number of training samples from the novel classes.

In this stage, all parameters of the model trained in stage 1 are frozen.Under the K-shot setting, each novel class contains K training samples. Given M novel classes, the total number of training samples for novel classes is $M \times K$. For the i-th novel class, all its training samples are first fed into the frozen feature extraction module. The extracted features are then passed through the base-class classifier to obtain the corresponding logits, followed by the Softmax function to produce $P_i^N \in \mathbb{R}^{K\times N}$. Here, each row of P_i^N represents the probability distribution of a training sample over the N base classes. Subsequently, the probabilities are averaged across the K samples to obtain $\bar{P}_i^N \in \mathbb{R}^{1\times N}$. The resulting $\bar{P}_i^N$ is regarded as the similarity vector between the i-th novel class and all base classes. Repeating this procedure for all novel classes yields the similarity matrix:

$$P^N = [\bar{P}_1^N, \bar{P}_2^N, \dots, \bar{P}_M^N] \tag{5}$$

Since each $\bar{P}_i^N$ characterizes the similarity between a novel class and all base classes, P_N can be regarded as the similarity matrix between the novel classes and the base classes.

Under the premise of sharing the same feature extraction module, the classification weights of the base classes can be regarded as discriminative representations of the corresponding classes. Therefore, the discriminative knowledge learned from the base classes can be transferred to the novel classes through the similarity matrix. The obtained similarity matrix P_N is then multiplied by the base-class classification weights W_b to generate the classification weights of the novel classes W_n. Finally, W_n is concatenated with W_b to obtain the joint classification weights W.

3.4. Fine-tuning

In this stage, the classification layer obtained in stage 1 is extended to a joint classification task over all classes. The classification weights are initialized using the merged joint classification weights obtained in stage 2. At this point, the model consists of the trained feature extraction module and the joint classifier.

During the fine-tuning stage, the parameters of the feature extraction module are frozen, and only the parameters of the classification module are updated. Under the K-shot setting, Ksamples are randomly selected from each base class and each novel class. Thus, the total number of training samples is $N_{bt} = (N + M) \times K$, and several epochs of fine-tuning are performed. This strategy is employed to alleviate class imbalance and mitigate overfitting caused by uneven sample distributions across classes. The cross-entropy loss function is adopted as the optimization objective, which is defined as:

$$\mathcal{L}\text{all} = -\frac{1}{N_{bt}} \sum_{i=1}^{N_{bt}} \sum_{j=1}^{N+M} (\mathbb{I}\,(y_i = c_j)\,\log(\hat{y}_{i,j})) \tag{6}$$

The objective is to minimize the loss function. Here,$\hat{y}_{i,j}$denotes the predicted

probability that the i-th sample belongs to class c_j, obtained via the Softmax function. $\mathbb{I}(\cdot)$is an indicator function,which equals 1 when the ground-truth label y_i of the i-th sample is c_j, and 0 otherwise.

4. Experiments

4.1. Dataset

HARTH dataset consists of 12 daily activity classes collected from 22 participants. The sampling frequency is 50 Hz, and each participant wears two triaxial accelerometers. In this study, data from both accelerometers are used. The window size is set to 1 s, and the sliding window step is set to 0.5 s.

HAR70+ dataset consists of 7 daily activity classes collected from 18 moderately to frail older adults. Each participant wears two triaxial accelerometers with a sampling frequency of 50 Hz. All accelerometer data are used in this study. The window size is set to 5 s, and the interval between adjacent windows is 5 s.

selfBACK dataset consists of 9 activity classes collected from 33 participants, with each activity recorded using two triaxial accelerometers. The sampling frequency is 100 Hz. In this study, data from both accelerometers are used. The window size is set to 5 s, and the stride between adjacent windows is 5 s.

4.2. Experiments Setup

During the first-stage training process, the learning rate is set to 0.001, and the Adam optimizer together with the cross-entropy loss function is adopted. In the few-shot learning experiments, 1-shot, 5-shot, and 10-shot settings are evaluated for each novel class. The classification weights of novel classes are concatenated with the original base-class classification weights. Support samples from both novel and base classes are used to fine-tune only the classifier parameters, using the Adam optimizer with a learning rate of 0.001 and the cross-entropy loss function.

4.3. Evaluation Metrics

We adopt the harmonic mean (HM) of base and novel classes as the evaluation metric. The harmonic mean is computed from the average class accuracy of base classes and the average class accuracy of novel classes. A higher HM value indicates that both the average class accuracy of base classes and the average class accuracy of novel classes are simultaneously high. During evaluation, HM is used as the primary metric because it provides a balanced measure of the model's performance on both base and novel classes.

4.4. Comparison With Baseline Methods

The proposed method was compared with five state-of-the-art baseline methods.

Classifier for Generalized Few-Shot Learning (CLS): This method achieves generalized human activity recognition by constructing three different classifiers.

Specifically, each sample is first evaluated by a base/novel discriminator to determine whether it belongs to base classes or novel classes. If a sample is classified as a base-class sample, it is further recognized by the base-class classifier; otherwise, it is classified using the novel-class classifier.

Few-Shot Category Weight Generator (FCWG): This method designs a few-shot weight generator for generalized few-shot human activity recognition. During training, part of the base-class data is used to train the weight generator. Then, support samples from novel classes are used to generate classification weights for novel classes, which are concatenated with the base-class classification weights for final prediction.

Deep Learning Model With Fine-Tuning (DFGM): This method first pre-trains a feature extraction network using base-class training samples, and adopts cosine similarity for classification. During fine-tuning, samples from novel classes are combined with downsampled base-class samples to construct a balanced training set, and the model is further optimized on this set.

Relational Network (RN): To address the problem that some classes contain only a few samples, this method proposes an improved relational network combined with metric learning and a multi-head attention mechanism. The model is capable of learning similarity relationships among different activity classes. In this framework, each query set contains only one query sample, while the support set consists of one sample from each class.

Autoencoder Framework (AE): This method employs an autoencoder-based architecture that projects features into a low-dimensional space and reconstructs them back to the original space. Classification is performed by measuring reconstruction similarity. To adapt it to the generalized few-shot setting, class-specific autoencoders are constructed for each activity class. During testing, features are fed into all autoencoders, and the class with the highest reconstruction similarity is selected.

4.5. Results Analysis

Table 1. HARTH Results.

Methods	HM(%)			
		1-shot	5-shot	10-shot
1	CLS	33.21	48.56	54.87
2	FCWG	27.93	31.01	37.13
3	DFGM	42.26	46.95	59.12
4	RN	36.72	44.87	50.32
5	AE	11.22	26.52	58.43
6	Ours	43.57	50.62	63.15

Table 2. HAR70+ Results.

Methods	HM(%)

		1-shot	5-shot	10-shot
1	CLS	35.65	45.84	47.14
2	FCWG	35.48	48.28	52.03
3	DFGM	34.78	51.88	52.71
4	RN	36.28	43.91	49.67
5	AE	15.06	28.74	44.42
6	Ours	40.43	60.37	64.37

Table 3. selfBACK Results.

Methods	HM(%)			
		1-shot	5-shot	10-shot
1	CLS	43.20	49.46	51.21
2	FCWG	54.65	57.82	61.47
3	DFGM	55.69	61.79	65.46
4	RN	31.46	47.40	48.55
5	AE	25.61	52.40	60.14
6	Ours	59.12	64.54	69.03

To evaluate the effectiveness of the proposed method, comparative experiments are conducted on three public datasets, namely HARTH, HAR70+, and selfBACK. Each experiment is repeated five times under 1-shot, 5-shot, and 10-shot settings.

On the HARTH dataset, the proposed method achieves HM values of 43.57%, 50.62%, and 63.15% under 1-shot, 5-shot, and 10-shot settings, respectively, outperforming all competing methods. In particular, under the 5-shot setting, the proposed method achieves 50.62%, which is approximately 2% higher than the second-best method CLS (48.56%). Under the 10-shot setting, it further improves to 63.15%, exceeding DFGM (59.12%) by about 4%. The detailed results are shown in Table 1.

On the HAR70+ dataset, the proposed method also achieved superior performance, with HM values of 40.43%, 60.37%, and 64.37% under 1-shot, 5-shot, and 10-shot settings, respec- tively. It achieves the best performance under both 5-shot and 10-shot settings. Compared with the strongest baseline, the improvement is approximately 8.49% and 11.66% under the 5-shot and 10-shot settings, respectively. These results demon- strate that as the number of novel-class samples increases, the proposed method can more effectively exploit novel-class information to refine classification weights, thereby improving overall recognition performance. The detailed results are shown in Table 2.

On the selfBACK dataset, the proposed method achieves HM values of 59.12%, 64.54%, and 69.03% under 1-shot, 5- shot, and 10-shot settings, respectively. It consistently outper- forms the best baseline methods by approximately 3%–4%. The detailed results are shown in Table 3.

5. Conclusion

To address the challenge of balancing base-class and novel-class recognition performance in generalized few-shot human activity recognition, this paper proposes a generalized few-shot human activity recognition model based on a Transformer architecture and a classification weight generation strategy, and conducts extensive experiments on multiple public datasets. The experimental results demonstrate that the proposed framework can effectively improve the recognition performance of novel classes while maintaining high recognition accuracy on base classes. Furthermore, under different few-shot settings, the proposed method consistently achieves superior overall performance compared with existing methods on three public datasets. The experimental results indicate that the proposed model can effectively alleviate the performance imbalance between base and novel classes in generalized few-shot human activity recognition, providing an effective solution for human activity recognition under limited labeled samples and offering valuable insights for further research in this field.

References

[1] Gu, F., Chung, M.-H., Chignell, M., Valaee, S., Zhou, B. and Liu, X. (2021) A Survey on Deep Learning for Human Activity Recognition. ACM Computing Surveys, 54, 1-34.

[2] Lara, O.D. and Labrador, M.A. (2013) A Survey on Human Activity Recognition Using Wearable Sensors. IEEE Communications Surveys & Tutorials, 15, 1192-1209.

[3] Su, X., Tong, H. and Ji, P. (2014) Activity Recognition with Smartphone Sensors. Tsinghua Science and Technology, 19, 235-249.

[4] Wang, Y., Yao, Q., Kwok, J.T. and Ni, L.M. (2020) Generalizing from a Few Examples: A Survey on Few-Shot Learning. ACM Computing Surveys, 53, 1-34.

[5] Gong, S., Shi, H., Yan, X., Fang, Y., Paul, A., Wu, Z. and Long, W. (2023) Human Activity Recognition with FMCW Radar Using Few-Shot Learning. IEEE Sensors Journal, 23, 17815-17824.

[6] Ding, X., Jiang, T., Zhong, Y., Wu, S., Yang, J. and Xue, W. (2021) Improving WiFi-Based Human Activity Recognition with Adaptive Initial State via One-Shot Learning. Proceedings of the IEEE Wireless Communications and Networking Conference, 1-6.

[7] Shi, Z., Zhang, J.A., Xu, R.Y. and Cheng, Q. (2022) Environment-Robust Device-Free Human Activity Recognition with Channel-State-Information Enhancement and One-Shot Learning. IEEE Transactions on Mobile Computing, 21, 540-554.

[8] Hu, P., Tang, C., Yin, K. and Zhang, X. (2021) WiGR: A Practical Wi-Fi-Based Gesture Recognition System with a Lightweight Few-Shot Network. Applied Sciences, 11, 3329.

[9] Li, X., He, Y., Zhang, J.A. and Jing, X. (2021) Supervised Domain Adaptation for Few-Shot Radar-Based Human Activity Recognition. IEEE Sensors Journal, 21, 25880-25890.

[10] Kasnesis, P., Chatzigeorgiou, C., Patrikakis, C.Z. and Rangoussi, M. (2021) Modality-Wise Relational Reasoning for One-Shot Sensor-Based Activity Recognition. Pattern Recognition Letters, 146, 90-99.

[11] Al Hafiz Khan, M.A. and Roy, N. (2022) Cross-Domain Unseen Activity Recognition Using Transfer Learning. Proceedings of the IEEE Annual Computer Software and Applications Conference, 684-693.

[12] Liu, Z., Wu, C. and Ye, W. (2022) Category-Extensible Human Activity Recognition Based on Doppler Radar by Few-Shot Learning. IEEE Sensors Journal, 22, 21952-21960.

[13] Wijayasingha, L. and Stankovic, J.A. (2022) Generalized Few-Shot Learning for Wearable Sensor-Based Human Activity Recognition. Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops, 328-334.

[14] Nguyen, L.T., Zeng, M., Tague, P. and Zhang, J. (2015) Recognizing New Activities with Limited Training Data. Proceedings of the ACM International Symposium on Wearable Computers, 67-74.

[15] Ghorpade, M., Chen, H., Liu, Y. and Jiang, Z. (2020) SMART: Emerging Activity Recognition with Limited Data for Multi-Modal Wearable Sensing. Proceedings of the IEEE International Conference on Big Data, 1316-1321.