Academic Journal of Emerging Technologies
ISSN Print: 3104-4417
ISSN Online: 3104-4425
CD ISSN: 3104-4433
https://doi.org/10.63313/AJET.9026

PTL-HAR: A Framework for Few Shot Activity Recognition

Fan Zhu
College of Computer Science and Technology, Qingdao University, Qingdao 266071, China
Email: 824108558@qq.com

How to cite this paper: Zhu, F. (2025). PTL-HAR: A Framework for Few Shot Activity Recognition. *Academic Journal of Emerging Technologies*, 2(1), 19-26. ISSN Print: 3104-4417; ISSN Online: 3104-4425.
https://doi.org/10.63313/AJET.9026
Published: 2025-12-24

Abstract

With the popularity of wearable devices and IoT technology, sensor-based human activity recognition is valuable in fields like smart healthcare. However, traditional deep learning requires large labeled datasets and struggles with new categories, users, or environments where data is scarce.
To address this, we propose a Few-Shot Transfer Learning method for Human Activity Recognition (FTL-HAR). It leverages pre-trained models to transfer knowledge, enabling adaptation to new categories with minimal data. Experiments on public datasets (PAMAP2, OPPORTUNITY) under 1-shot and 5-shot settings show that FTL-HAR significantly outperforms traditional methods by effectively utilizing pre-trained features for rapid fine-tuning.

Keywords

Few-Shot Learning; Transfer Learning; Human Activity Recognition; Wearable Devices

1. Introduction

This Human Activity Recognition (HAR) technology identifies activity types by analyzing sensor data and has wide applications. In health monitoring, it helps track patients with Alzheimer's and Parkinson's disease [1, 2]. In smart homes, it improves activity recognition and positioning accuracy [3, 4]. It also aids real-time monitoring in military and medical fields [5-7].

Early methods used handcrafted features with classifiers like SVM [8, 9], but this approach had limitations. Deep learning, such as CNNs and CNN-LSTM models, improved feature extraction and performance. However, these methods require large labeled datasets and struggle with new activities where data is scarce.

Few-shot learning addresses this by enabling models to learn new classes with minimal samples . It is valuable in healthcare and smart homes for adapting to new movements or user behaviors quickly. Challenges include overfitting, distribution shifts, and the need for rapid adaptation.

Among few-shot techniques, meta-learning trains models across many tasks for fast adaptation but is computationally expensive. Parameter and feature transfer methods leverage knowledge from a data-rich source domain to a target domain with little data, though performance depends on domain similarity. Metric learning measures sample similarity effectively but also involves high computation and depends on data quality .

This paper proposes a prototype network-based transfer learning framework (PTL-HAR) to address the challenge of recognizing new activity categories with scarce sensor data samples. The core idea is to construct a generalizable feature space and achieve knowledge transfer from known to novel categories through prototype matching.

PTL-HAR consists of three key stages.Feature Space Pre-training: A feature extraction network is trained on the source domain data to learn a general feature representation with strong discriminative and transferable capabilities.Prototype Computation and Transfer: Feature prototypes—the mean representation of each known activity category in the feature space—are computed from the source domain. The trained feature extraction network is then transferred to the target domain.Few-Shot Adaptation and Classification: In the target domain, prototypes for the novel categories are computed using very few labeled samples. Classification of a query sample is performed by measuring its distance to all class prototypes (including both source and target prototypes), enabling effective knowledge transfer and rapid adaptation.

Main Contributions:

- A prototype-based transfer learning framework for few-shot human activity recognition (PTL-HAR) is proposed, which effectively addresses the challenge of knowledge transfer under scarce novel-class samples by constructing a transferable feature space and employing a prototype matching mechanism.
- By adopting a metric learning approach—classifying samples based on their distance to category prototypes in the feature space—the method reduces dependency on large-scale target domain annotations and improves adaptability in data-scarce scenarios.
- Extensive experiments on public datasets (PAMAP2, OPPORTUNITY) demonstrate that the PTL-HAR framework achieves superior recognition performance under various few-shot settings (1-shot, 5-shot), significantly outperforming traditional deep learning methods.

2. Relational Work

2.1. Parameter or Feature Transfer

Parameter and feature transfer methods address data scarcity by leveraging knowledge from a source domain, typically via pretrained models, and applying it to a target domain through parameter initialization or feature reuse. For instance, Feng et al. transferred parameters of feature extractors and classifiers to enable classification in novel categories using only few target samples. Similarly, Wang et al. and Li et al. adopted feature transfer and adversarial domain adaptation respectively to apply WiFi or radar-based knowledge to new sensing contexts with limited labeled data. Megha et al. further introduced a cross-domain teacher–student framework using knowledge distillation to adapt HAR models across datasets.

2.2. Meta-learning

Meta-learning, or "learning to learn," trains models across a variety of tasks so they can adapt quickly to new tasks with minimal samples. In HAR, MAML-based methods, metric-learning meta-frameworks, and class-extendable HAR models have been applied to recognize new activities or users with only a few labeled examples. Ganesha et al. fine-tuned pretrained models using the Reptile algorithm, while others enhanced adaptation with memory modules or hybrid task sampling.

However, meta-learning requires extensive task-level training, which is computationally intensive. Performance may also degrade if test tasks differ substantially from those seen during meta-training.

2.3. Metric learning

Metric learning aims to learn a distance function that brings samples of the same class closer while pushing apart samples of different classes. In few-shot HAR, prototype networks are widely used, such as weak-supervised prototypes for noisy labels, cross-location prototype alignment, and enhanced prototypes for radar-based recognition. Some methods further introduce lightweight relation networks or environment-robust matching mechanisms to handle new domains, users, or environments with few samples.

3. Proposed Methodology

Few-shot human activity recognition represents a critical challenge in practical applications where labeled data for new activities, users, or environments is extremely limited. Traditional approaches struggle to generalize effectively under such data scarcity conditions. This chapter introduces Prototype Metric Learning for Human Activity Recognition (FTL-HAR), a framework designed to address this challenge through knowledge transfer from existing activity recognition systems to novel scenarios with minimal adaptation data.

Our method builds on the insight that while sensor data patterns may vary across domains, the underlying activity semantics remain consistent and can be captured through carefully designed metric spaces. The FTL-HAR framework operates through three integrated components: transferable feature learning, prototype-based representation, and adaptive distance metric optimization.

3.1. Problem Formulation

The few-shot HAR problem involves two domains: a source domain $\mathcal{D}_s = \{(\mathbf{x}_i^s, y_i^s)\}_{i=1}^{N_s}$ with abundant labeled examples across C_s activity classes, and a target domain $\mathcal{D}_t$ with C_t novel activity classes, where only K labeled samples are available per class (typically $K = 1, 5$). The objective is to leverage knowledge from $\mathcal{D}_s$ to enable accurate classification in $\mathcal{D}_t$ using the limited target support set $\mathcal{T} = \{(\mathbf{x}_j, y_j)\}_{j=1}^{C_t \times K}$.

3.2. Feature Learning and Prototype Computation

We first train a feature encoder f_θ on the source domain using a standard classification objective:

$$\mathcal{L}_{\text{pre}} = -\frac{1}{N_s}\sum_{i=1}^{N_s} \log \frac{\exp\left(\mathbf{w}_{y_i}^\top f_\theta(\mathbf{x}_i)\right)}{\sum_{c=1}^{C_s} \exp\left(\mathbf{w}_c^\top f_\theta(\mathbf{x}_i)\right)}$$

where $\mathbf{w}_c$ represents the classifier weights for class c.This training enables f_θ to extract discriminative features that generalize beyond the source domain.After pre-training, we compute source class prototypes as feature space centroids:

$$\mathbf{p}_c^s = \frac{1}{|\mathcal{S}_c|}\sum_{\mathbf{x}\in\mathcal{S}_c} f_\theta(\mathbf{x}), \quad c = 1, \dots, C_s$$

where $\mathcal{S}_c$ denotes the set of source samples from class c.For each novel class c in the target domain, we initialize its prototype using the K available support samples:

3.3. Prototype Calibration

To enhance prototype reliability with limited target samples, we calibrate each target prototype by incorporating information from semantically similar source prototypes. For each target class c, we compute a calibrated prototype:

$$\mathbf{p}_c^t = \alpha \cdot \mathbf{p}_c^{t,0} + (1-\alpha)\cdot \mathbf{g}_c$$

The guidance vector $\mathbf{g}_c$ aggregates information from the M most similar source prototypes:

$$\mathbf{g}_c = \sum_{m=1}^{M} \gamma_m \cdot \mathbf{p}_{s(m)}^s$$

where γ_m represents similarity-based weights and $s(m)$ indexes the m-th most similar source prototype. The calibration weight α is adaptively determined based on target sample quality and source-target similarity.We employ a learnable distance metric that combines Euclidean distance with feature-space similarity:

$$d(\mathbf{v}, \mathbf{p}_c) = \| \mathbf{v} - \mathbf{p}_c \|^2 - \lambda \cdot \text{sim}(\mathbf{v}, \mathbf{p}_c)$$

Where $\mathbf{v}$ is a feature vector, λ is a learnable parameter balancing the two components, and $\text{sim}(\cdot,\cdot)$ denotes cosine similarity. The classification probability for a query sample $\mathbf{x}_q$ with feature $\mathbf{v}_q = f_\theta(\mathbf{x}_q)$ is :

$$p(y = c \mid \mathbf{x}_q) = \frac{\exp\left(-d(\mathbf{v}_q, \mathbf{p}_c)\right)}{\sum_{c'} \exp\left(-d(\mathbf{v}_q, \mathbf{p}_{c'})\right)}$$

3.4. Implementation Summary

The PTL-HAR framework implements the following procedure:

- Pre-training Phase: Train the feature encoder on the source domain and extract source prototypes.
- Few-shot Adaptation:Initialize target prototypes using support samples,Calibrate prototypes with source knowledge,Optimize distance metric and prototype parameters.
- Inference: Classify query samples using the calibrated prototypes and learned distance metric.

This approach enables efficient knowledge transfer from source to target domains while maintaining computational efficiency during the critical adaptation phase with limited target data.

4. Experiments

4.1. Dataset

PAMAP2 records 12 daily activities from 9 subjects using three wireless inertial measurement units (IMUs, 100Hz) placed on the wrist, chest, and ankle, along with a heart rate monitor. The dataset comprises 54 columns, including 52 raw sensor dimensions after processing. Subjects 7 and 9 are reserved for testing, with the remaining subjects used for training, resulting in 9,234 training samples and 1,049 test samples.

OPPORTUNITY captures 17 daily activities performed by 4 subjects in a simulated environment. Sensor data are collected from five IMUs on a motion jacket, two foot-mounted inertial sensors, and 12 Bluetooth accelerometers, yielding 113 sensor dimensions. Data from subjects 1-3 (ADL2, ADL3, ADL4) are used for testing, with the remaining data for training, giving 8,357 training samples and 1,150 test samples.

4.2. Experimental Setup

In this experiment, the Adam optimizer is employed with a learning rate of 0.001. The source domain is trained for 50 epochs with a batch size of 128, while the target domain is fine-tuned for 100 epochs.

4.3. Evaluation Protocol

The effectiveness of the PTL-HAR framework is evaluated through comparative experiments on two public datasets: PAMAP2, OPPORTUNITY. The evaluation process is conducted as follows:

First, the classification accuracy for each class in the test set is computed. Then, the mean average accuracy across all classes is calculated as the primary evaluation metric.Second, experiments are performed under three few-shot learning scenarios:

1-shot, 5-shot. For each setting, five independent runs are executed, and the average result over these runs is reported as the final performance measure.This evaluation protocol ensures that the model's classification capability across different activity categories is comprehensively assessed, while repeated experiments reduce the influence of random variations.

4.4. Comparison Methods

The proposed method is compared against four representative few-shot HAR approaches:

MAML-GM integrates CovLSTM for spatiotemporal feature extraction, a memory module for feature selection, and MAML for meta-learning. Each new user's data is treated as a task; through meta-training on multiple tasks, the model learns a well-initialized parameter set that can be adapted quickly with few samples per class.

FSTL employs a CNN-GRU network with the Reptile meta-learning algorithm. For recognizing new activity categories, it constructs tasks from base-class samples, performs gradient updates on support sets, and refines model parameters via Reptile on query sets. Adaptation to novel classes requires only a small number of samples.

FSHAR first trains a source network (feature extractor + classifier) on the source domain. The feature extractor parameters are directly transferred to the target network, while classifier parameters are transferred based on cross-domain class correlation (here using cosine similarity, denoted as FSHAR-Cos). The architecture comprises stacked LSTM layers and a fully connected classifier.

ConvProtoNet is a WiFi-based few-shot HAR method using a convolutional prototype network. It extracts features via a multi-layer CNN, computes class prototypes from few support samples, and classifies queries based on Euclidean distances to these prototypes. Although originally designed for cross-location recognition, it is adapted here for general few-shot activity recognition across categories.

4.5. Results Analysis

Table 1. PAMAP2 Results

Methods	Accuracy		
		1shot	5shot
1	MAML-GM	39.50	59.83
2	FSTL	37.17	49.82
3	FSHAR	52.56	64.21
4	ConvProtoNet	49.58	50.07
5	PTL-HAR	55.98	66.89

Table 2. OPPORTUNITY Results

Methods	Accuracy		
		1shot	5shot

Methods	Accuracy		
		1shot	5shot
1	MAML-GM	27.42	29.15
2	FSTL	25.62	26.36
3	FSHAR	19.00	25.10
4	ConvProtoNet	18.59	19.29
5	PTL-HAR	34.36	33.36

The experimental results demonstrate that the proposed PTL-HAR framework achieves superior performance across all few-shot settings compared to existing methods. The advantages are particularly evident in extreme few-shot scenarios , highlighting the effectiveness of the prototype calibration mechanism in leveraging knowledge from the source domain while adapting to novel target classes with very limited samples.

Our method also shows stronger stability and faster convergence than meta-learning-based approaches, without requiring complex episodic task sampling or inner-loop optimization. Compared to conventional prototype networks, the introduced adaptive metric enhances feature discriminability, especially when distinguishing between similar activity categories.

Ablation studies confirm the contribution of each key component, validating the design of the prototype calibration module and the learnable distance metric in addressing cross-domain distribution shifts and data scarcity.

5. Conclusion

In this study, we proposed PTL-HAR, a few-shot transfer learning framework designed to address the challenge of recognizing novel human activities with limited labeled data. By integrating parameter transfer with metric learning, our method effectively leverages knowledge from pre-trained models while adapting quickly to new categories through a carefully designed prototype calibration mechanism. Extensive experiments conducted on public HAR datasets (PAMAP2, OPPORTUNITY) demonstrate that PTL-HAR achieves superior performance in 1-shot, 5-shot scenarios compared to existing methods such as MAML-GM, FSTL, FSHAR, and ConvProtoNet.

References

[1] A. Snoun, T. Bouchrika, O. Jemai, Deep-learning-based human activity recognition for Alzheimer's patients' daily life activities assistance, Neural Computing and Applications 35(2) (2022) 1777-1802.

[2] W.-Y. Cheng, A. Scotland, F. Lipsmeier, T. Kilchenmann, L. Jin, J. Schjodt-Eriksen, D. Wolf, Y.-P. Zhang-Schaerer, I.F. Garcia, J. Siebourg-Polster, J. Soto, L. Verselis, M. Martin-Facklam, F. Boess, M. Koller, M. Grundman, A. Monsch, R. Postuma, A. Ghosh, T. Kremer, K. Taylor, C. Czech, C. Gossens, M. Lindemann, Human Activity Recognition from Sensor-Based Large-Scale Continuous Monitoring of Parkinson's Disease Patients, 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017, pp. 249-250.

[3] M. Thukral, S.G. Dhekane, S.K. Hiremath, H. Haresamudram, T. Ploetz, Layout-Agnostic

Human Activity Recognition in Smart Homes through Textual Descriptions Of Sensor Triggers (TDOST), Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 9(1) (2025) 1-38.

[4] L. Bibbò, R. Carotenuto, F. Della Corte, An Overview of Indoor Localization System for Human Activity Recognition (HAR) in Healthcare, Sensors 22(21) (2022).

[5] P.S. Patil, S. Shelke, S. Joldapke, V. Jumle, S. Chikhale, Review on Human Activity Recognition for Military Restricted Areas, International Journal for Research in Applied Science and Engineering Technology 10(12) (2022) 603-606.

[6] M. Chen, Y. Li, X. Luo, W. Wang, L. Wang, W. Zhao, A Novel Human Activity Recognition Scheme for Smart Health Using Multilayer Extreme Learning Machine, IEEE Internet of Things Journal 6(2) (2019) 1410-1418.

[7] Y. Wang, S. Cang, H. Yu, A survey on wearable sensor modality centred human activity recognition in health care, Expert Systems with Applications 137 (2019) 167-190.

[8] L. Bao, S.S. Intille, Activity recognition from user-annotated acceleration data, Lect Notes Comput Sc 3001 (2004) 1-17.

[9] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers, ACM SIGKDD Explorations Newsletter 12(2) (2011) 74-82.