DuLPA: Dual-Level Prototype Alignment for Unsupervised Domain Adaptation in Activity Recognition from Wearables

Kishor Kumar Bhaumik
University of California, Riverside
kbhau001@ucr.edu

Jia Chen
University of California, Riverside
jiac@ucr.edu

Evangelos E. Papalexakis
University of California, Riverside
epapalex@cs.ucr.edu

Abstract

In wearable human activity recognition (WHAR), models often falter on unseen users due to behavioral and sensor differences.Without target labels, unsupervised domain adaptation (UDA) can help improve cross-user generalization. However, many WHAR UDA methods either pool all source users together or perform one-to-one source–target alignment, ignoring individual differences and risking negative transfer. To address this critical limitation, we propose ***DuLPA***—**Du**al-**L**evel **P**rototype **A**lignment method for unsupervised cross-user domain adaptation. First, it aligns class prototypes between each source user and the target to capture individual variation; a convex reweighting further handles class imbalance. Second, a BLUP-based fusion forms robust global class prototypes by optimally weighting domain-specific ones using estimated within- and between-domain variances. On four public datasets, ***DuLPA*** outperforms several baselines, improving macro-F1 by 5.34%. Our source code is available at https://github.com/Kishor-Bhaumik/DuLPA-TS4H.

1 Introduction

Human Activity Recognition (HAR) is widely applied in healthcare and manufacturing, powered by wearable devices with IMU sensors [1, 2]. Despite strong deep learning results [3, 4], cross-user wearable HAR (WHAR) remains difficult due to user-specific distribution shifts [5]. Supervised adaptation [6] requires costly labels, motivating Unsupervised Domain Adaptation (UDA). Existing WHAR UDA methods either pool sources into one domain [7] or use one-to-one alignment [8], ignoring user diversity and risking negative transfer. Multi-source approaches [9] and prototype-based methods [10] show promise but often average across domains and neglect reliability, which limits robustness under label imbalance [8]. Beyond WHAR, prior work explores explainability [11], fairness [12], and cross-modal transfer [13], but prototype-based UDA for WHAR remains underexplored. In light of this, we propose ***DuLPA***, a **Du**al-**L**evel **P**rototype **A**lignment framework. At the base level, DuLPA aligns each source with the target using convex reweighting for label shift. At the upper level, a BLUP-inspired fusion builds global prototypes by weighting sources by reliability. A bidirectional prototype alignment loss further enforces semantic consistency. Our contributions are as follows, **(1)** Dual-level alignment capturing user-specific variations and fusing sources by reliability. **(2)** Adaptive convex reweighting to address class imbalance. **(3)** Bidirectional prototype alignment loss for robust transfer. **(4)** Comprehensive results on four HAR datasets, with up to 5.34% macro-F1 gain.

Workshop on Learning from Time Series for Health, 39th Conference on Neural Information Processing Systems (NeurIPS 2025).

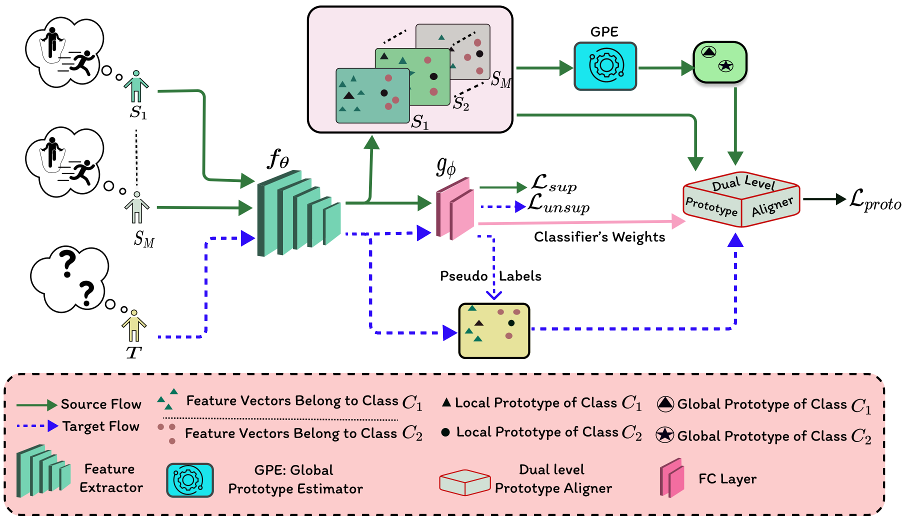


Figure 1: Framework of ***DuLPA***. Given M labeled sources and an unlabeled target T, features are extracted by f_θ. Domain prototypes are fused via the BLUP-based GPE into global prototypes μ^G, weighted by within- and between-domain variance, which then align bidirectionally with target features.

2 Problem Setup

We study cross-user HAR via unsupervised domain adaptation. Given M labeled source domains $\{S_1, \ldots, S_M\}$ with samples $D_{S_i} = \{(x_j^{S_i}, y_j^{S_i})\}$ and an unlabeled target domain T with data $D_T = \{x_j^T\}$, all domains share the same label space $\{1, \ldots, C\}$ but differ in distributions and class frequencies (label shift). Let $f_\theta : \mathcal{X} \to \mathbb{R}^d$ be the feature extractor and $g_\phi : \mathbb{R}^d \to \{1, \ldots, C\}$ the classifier (Figure 1). The goal is to learn f_θ, g_ϕ that generalize to the unlabeled target T.

3 Methodology

Prototype-to-Prototype Alignment. For each activity class c in each source domain S_i, *we compute a class prototype* $\mu_c^{S_i}$ *as the average feature representation of all samples from class c.* Similarly, target pseudo-prototypes μ_c^T are constructed using pseudo-labels. We align source and target prototypes by minimizing their distances:

$$\mathcal{L}_{\text{direct}} = \sum_{i=1}^{M} \sum_{c=1}^{C} \alpha_c^{(i)} \, \|\mu_c^{S_i} - \mu_c^T\|_2^2. \tag{1}$$

To account for label-shift between users, we adopt a BBSE-guided convex reweighting strateg[14]. We estimate class-level reweighting coefficients $\alpha^{(i)} \in \mathbb{R}^C$ via the following constrained least-squares problem:

$$\alpha^{(i)} = \arg\min_{\alpha \geq 0} \|\alpha^\top C^{(i)} - \hat{p}_T\|_2^2, \quad \text{s.t. } (\alpha^{(i)})^\top p_{S_i} = 1.$$

where $C^{(i)}$ is the confusion matrix of a classifier trained on S_i and $\hat{p}_T$ is the empirical label distribution of the target.This allows us to emphasize classes that are more relevant to the target user and down-weight overrepresented source classes, yielding more robust prototype alignment.

Prototype-to-Feature Alignment. Direct source–target alignment is necessary but insufficient: user-to-user variability in physiology, movement, and device placement injects domain-specific noise and inconsistent features, so equal weighting ignores that some users provide more reliable and relevant patterns than others. To overcome such limitations, we propose a BLUP (Best Linear Unbiased Prediction) based fusion approach [15, 16] that allocates more weight to source domains with more reliable (low-variance) class prototypes, as determined by both within- and between-domain variability.

*Global Prototype Estimation (GPE).*We model each domain prototype z_i^c as a noisy observation of a latent global prototype θ^c:

$$z_i^c = \theta^c + b_i^c + \epsilon_i^c,$$

where $b_i^c \sim \mathcal{N}(0, \tau_c^2 I_d)$ captures between-domain and $\epsilon_i^c \sim \mathcal{N}(0, \sigma_{i,c}^2 I_d)$ captures within-domain variations. The variances are estimated as,

$$\sigma_{i,c}^2 = \frac{1}{n_{i,c}-1} \sum_{x \in D_{S_i}^c} \|f_\theta(x) - z_i^c\|^2, \quad \tau_c^2 = \max\left(0, \frac{1}{k_c - 1} \sum_{i \in \mathcal{D}_c} \|z_i^c - \bar{z}^c\|^2\right).$$

The global prototype is obtained via inverse-variance weighting:

$$w_{i,c} = \frac{1}{\sigma_{i,c}^2 + \tau_c^2}, \quad \mu_c^G = \frac{\sum_{i \in \mathcal{D}_c} w_{i,c} z_i^c}{\sum_{i \in \mathcal{D}_c} w_{i,c}}.$$

For large τ_c^2, we interpolate with classifier weights ϕ_c:

$$\gamma_c = \frac{1}{1 + \tau_c^2}, \quad \mu_c^G \leftarrow \gamma_c \mu_c^G + (1 - \gamma_c)\phi_c.$$

Bidirectional Prototype Alignment. Inspired form PCT [10], we align target features and global prototypes with *bidirectional* soft assignments: target→prototype and prototype→target. Let $c(\mu, \mathbf{f}) = 1 - \frac{\mu \cdot \mathbf{f}}{\|\mu\| \, \|\mathbf{f}\|}$ be the cosine dissimilarity. Using class-prior weights $p(\mu_c^G)$ and softmax assignments $\pi(\mu_c^G \,|\, \mathbf{f}_j^T)$ (over classes) and $\pi(\mathbf{f}_j^T \,|\, \mu_c^G)$ (over batch) as in PCT, the bidirectional transport loss is

$$\mathcal{L}_{\text{bp}} = \mathbb{E}_{\{\mathbf{f}_j^T\}} \Big[\sum_{c=1}^{C} \pi(\mu_c^G \,|\, \mathbf{f}_j^T)\, c(\mu_c^G, \mathbf{f}_j^T) \; + \; \sum_{c=1}^{C} p(\mu_c^G) \sum_{j=1}^{B} \pi(\mathbf{f}_j^T \,|\, \mu_c^G)\, c(\mu_c^G, \mathbf{f}_j^T) \Big], \tag{2}$$

which discourages collapse and promotes balanced class coverage. We estimate $p(\mu_c^G)$ with a lightweight EM update from posteriors on the target batch:

$$p^{(t+1)}(\mu_c^G) = (1 - \beta^{(t)})\, p^{(t)}(\mu_c^G) + \beta^{(t)} \Big(\frac{1}{B} \sum_{j=1}^{B} \pi^{(t)}(\mu_c^G \,|\, \mathbf{f}_j^T) \Big),$$

using a decaying $\beta^{(t)}$; for balanced targets we keep $p(\mu_c^G)$ uniform. Combining with direct source–target prototype alignment (Eq. 1 & Eq. 2), our prototype objective is

$$\mathcal{L}_{\text{proto}} = \mathcal{L}_{\text{direct}} + \mathcal{L}_{\text{bp}}. \tag{3}$$

Adversarial Domain Adaptation. We further incorporate adversarial learning to promote domain-invariant features. A feature extractor $f(\cdot)$, classifier $g(\cdot)$, and domain discriminator $D(\cdot)$ are trained jointly: D distinguishes source vs. target, while f is optimized to confuse D. The supervised and adversarial losses are:

$$\begin{aligned} \mathcal{L}_{\text{sup}} &= \mathbb{E}_{(x_i^s, y_i^s) \sim \mathcal{D}_s} \mathcal{L}_{\text{ce}}(g(f(x_i^s)), y_i^s) \\ \mathcal{L}_{\text{adv}} &= \sum_{k=1}^{M} \left(\mathbb{E}_{x_i^s \sim \mathcal{S}_k} \log[D(f(x_i^s))] + \mathbb{E}_{x_i^t \sim \mathcal{T}} \log[1 - D(f(x_i^t))] \right). \end{aligned} \tag{4}$$

We also adopt the Minimum Class Confusion (MCC) loss [17] to regularize unlabeled target predictions without relying on pseudo-labels:

$$\mathcal{L}_{\text{unsup}} = L_{\text{MCC}}(\hat{Y}_t), \quad \hat{Y}_t = g(f(X_t)). \tag{5}$$

The Final Objective. At the end of the approach, let us integrate all of these losses together, i.e, the prototype loss in Eq. 3, supervised classification loss $\mathcal{L}_{sup}$, domain adversarial loss $\mathcal{L}_{adv}$ described in Eq. 4 and the unsupervised loss for the unlabeled target domain in Eq. 5. Finally, we can obtain the final objective as follows:

$$\mathcal{L}_{total} = \lambda_1 \, \mathcal{L}_{sup} + \lambda_2 \mathcal{L}_{adv} + \lambda_3 \mathcal{L}_{unsup} + \lambda_4 \mathcal{L}_{proto}. \tag{6}$$

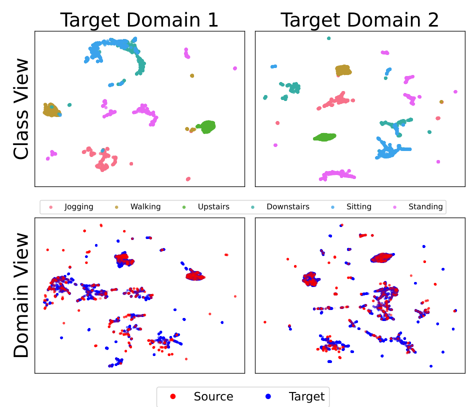
Here, λ's are the loss scaling coefficients.

Model	Opportunity	SBHAR	WISDM	PAMAP2
CoDATs (KDD'20)[22]	61.79±2.87	78.82±1.15	61.47±5.97	88.20±3.15
PCT (NeurIPS'21) [10]	68.86±4.09	<u>88.67±1.13</u>	68.67±6.16	90.67±1.94
DWLR (IJCAI'24)[8]	69.16	87.33	71.15	93.17
SWL-Adapt (AAAI'23)[7]	<u>70.31 ± 2.66</u>	85.62±0.97	72.98±4.82	<u>96.98±2.82</u>
μDAR (ICDM'24) [23]	66.25	82.97	<u>77.98</u>	95.14
***DuLPA* (ours)**	**75.65 ± 2.48**	**90.91± 0.81**	**79.74 ± 4.58**	**98.41± 1.78**

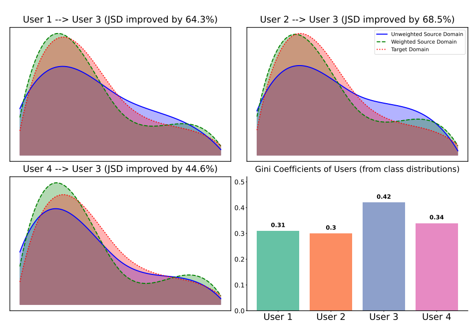
Table 1: Overall performance (macro-F1, in %) of ***DuLPA*** and baselines. The top value is highlighted in bold **blue** and the second best in **green** with underline. Standard deviations (±std in %) indicate result consistency.

Handling Class Imbalance with Convex Weighting. We assess the effectiveness of our convex weighting strategy in navigating class imbalance, as shown in Figure 3(b). Gini coefficients (0.31, 0.30, 0.42, 0.34 for Users 1–4) highlight varying imbalance levels, with User 3 being the most skewed. Jensen-Shannon Divergence (JSD) results confirm our method's robustness, achieving reductions of 64.3%, 68.5%, and 44.6% when adapting Users 1, 2, and 4 to User 3, respectively. The largest improvement (68.5%) occurs when bridging from the most balanced (User 2) to the most imbalanced (User 3), demonstrating dynamic calibration of source importance. Overall, convex weighting effectively addresses cross-user label imbalance in UDA.

4.1 Ablation Study. To assess DuLPA's core components, we analyze the impact of convex weighting and BLUP-based fusion (Table 2). The baseline encoder shows moderate performance, while convex weighting yields substantial gains (e.g., +8.58% on Opportunity) by mitigating label shifts. BLUP-based fusion also improves results, particularly on WISDM (+5.00%). Combining both delivers the best performance across all datasets; for instance, on SBHAR, their integration boosts macro-F1 by 6.24% over convex weighting alone, confirming their complementary benefits.



(a) SBHAR feature visualization



(b) Opportunity distribution alignment

Figure 3: Experimental results analysis.

4 Experiments

Result Analysis. We conduct experiments on four widely-used public HAR datasets: Opportunity [18],SBHAR [19],WISDM [20] and PAMAP2 [21]. We compare ***DuLPA*** against recent domain adaptation methods across four HAR datasets, with results summarized in Table 1. Overall, ***DuLPA*** consistently achieves the best performance, surpassing strong baselines such as prototype-based PCT, domain alignment approaches like DWLR and μDAR, and multi-source adaptation methods like CoDATs and SWL-Adapt. Notably, ***DuLPA*** delivers substantial improvements on challenging smartphone-based datasets, outperforming the closest competitor by +5.34% on Opportunity and +1.24% on WISDM. Figure 3(a) further illustrates that our approach yields well-separated activity clusters and strong source–target feature alignment.

Table 2: Component Analysis of ***DuLPA***.

Encoder Only	Convex Weighting	BLUP Fusion	Opportunity macro-F1	WISDM macro-F1	SBHAR macro-F1	PAMAP2 macro-F1
√	X	X	62.54	71.67	84.86	90.17
√	√	X	71.12	78.29	88.67	95.19
√	X	√	65.67	76.67	83.67	91.28
√	√	√	**75.65**	**79.74**	**90.91**	**97.14**

5 Conclusion

We proposed ***DuLPA***, an unsupervised domain adaptation framework with dual-level prototype learning for wearable based human activity recognition. At the base level, we align sources to the target using convex reweighting to handle class-prior shift, adaptively modulating each source's contribution. At the upper level, a BLUP-inspired fusion builds global prototypes by weighting sources via within- and between-domain variability, strengthening transfer. Extensive experiments on four benchmark datasets demonstrate ***DuLPA***'s superior performance and its effectiveness in cross-user adaptation for wearable human activity recognition. Our findings suggest that ***DuLPA*** offers a promising solution for personalized activity recognition in scenarios with unlabeled data for new users. Supplementary material is attached after the reference.

6 Acknowledgment

This research was supported in part by the National Science Foundation under grant no. No. 2431569 and CREST Center for Multidisciplinary Research Excellence in CyberPhysical Infrastructure Systems (MECIS) grant no. 2112650. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

References

[1] Sozo Inoue, Paula Lago, Tahera Hossain, Tittaya Mairittha, and Nattaya Mairittha. Integrating activity recognition and nursing care records: The system, deployment, and a verification study. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3):1–24, 2019.

[2] Kaixuan Chen, Dalin Zhang, Lina Yao, Bin Guo, Zhiwen Yu, and Yunhao Liu. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Computing Surveys (CSUR)*, 54(4):1–40, 2021.

[3] Hangwei Qian, Sinno Jialin Pan, Bingshui Da, and Chunyan Miao. A novel distribution-embedded neural network for sensor-based activity recognition. 2019.

[4] Xiyuan Zhang, Diyan Teng, Ranak Roy Chowdhury, Shuheng Li, Dezhi Hong, Rajesh Gupta, and Jingbo Shang. Unimts: Unified pre-training for motion time series. *Advances in Neural Information Processing Systems*, 37:107469–107493, 2024.

[5] Ling Chen, Yi Zhang, and Liangying Peng. Metier: a deep multi-task learning based activity and user recognition model using wearable sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–18, 2020.

[6] Seyed Ali Rokni, Marjan Nourollahi, and Hassan Ghasemzadeh. Personalized human activity recognition using convolutional neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[7] Rong Hu, Ling Chen, Shenghuan Miao, and Xing Tang. Swl-adapt: An unsupervised domain adaptation model with sample weight learning for cross-user wearable human activity recognition. In *Proceedings of the AAAI Conference on artificial intelligence*, volume 37, pages 6012–6020, 2023.

[8] Juren Li, Yang Yang, Youmin Chen, Jianfeng Zhang, Zeyu Lai, and Lujia Pan. Dwlr: domain adaptation under label shift for wearable sensor. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 4425–4433, 2024.

[9] Chuan-Xian Ren, Yong-Hui Liu, Xi-Wen Zhang, and Ke-Kun Huang. Multi-source unsupervised domain adaptation via pseudo target domain. *IEEE Transactions on Image Processing*, 31:2122–2135, 2022.

[10] Korawat Tanwisuth, Xinjie Fan, Huangjie Zheng, Shujian Zhang, Hao Zhang, Bo Chen, and Mingyuan Zhou. A prototype-oriented framework for unsupervised domain adaptation. *Advances in Neural Information Processing Systems*, 34:17194–17208, 2021.

[11] Jeya Vikranth Jeyakumar, Ankur Sarker, Luis Antonio Garcia, and Mani Srivastava. X-char: A concept-based explainable complex human activity recognition model. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 7(1):1–28, 2023.

[12] Youpeng Li, Xuyu Wang, and Lingling An. Hierarchical clustering-based personalized federated learning for robust and fair human activity recognition. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 7(1):1–38, 2023.

[13] Sejal Bhalla, Mayank Goel, and Rushil Khurana. Imu2doppler: Cross-modal domain adaptation for doppler-based activity recognition using imu data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–20, 2021.

[14] Zachary Lipton, Yu-Xiang Wang, and Alexander Smola. Detecting and correcting for label shift with black box predictors. In *International conference on machine learning*, pages 3122–3130. PMLR, 2018.

[15] Charles R Henderson. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, pages 423–447, 1975.

[16] Arthur S Goldberger. Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association*, 57(298):369–375, 1962.

[17] Ying Jin, Ximei Wang, Mingsheng Long, and Jianmin Wang. Minimum class confusion for versatile domain adaptation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16*, pages 464–480. Springer, 2020.

[18] Daniel Roggen, Alberto Calatroni, Mirco Rossi, Thomas Holleczek, Kilian Förster, Gerhard Tröster, Paul Lukowicz, David Bannach, Gerald Pirkl, Alois Ferscha, et al. Collecting complex activity datasets in highly rich networked sensor environments. In *2010 Seventh international conference on networked sensing systems (INSS)*, pages 233–240. IEEE, 2010.

[19] Jorge-L Reyes-Ortiz, Luca Oneto, Albert Samà, Xavier Parra, and Davide Anguita. Transition-aware human activity recognition using smartphones. *Neurocomputing*, 171:754–767, 2016.

[20] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82, 2011.

[21] Attila Reiss and Didier Stricker. Introducing a new benchmarked dataset for activity monitoring. In *2012 16th international symposium on wearable computers*, pages 108–109. IEEE, 2012.

[22] Garrett Wilson, Janardhan Rao Doppa, and Diane J Cook. Multi-source deep domain adaptation with weak supervision for time-series sensor data. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1768–1778, 2020.

[23] Indrajeet Ghosh, Garvit Chugh, Abu Zaher Md Faridee, and Nirmalya Roy. Unsupervised domain adaptation for action recognition via self-ensembling and conditional embedding alignment. *arXiv preprint arXiv:2410.17489*, 2024.