MixMatch-based Semi-supervised Learning Approach for Cross-domain Locomotion and Transportation Mode Recognition

Hao Wang*
Shenzhen Institutes of Advanced Technology, Chinese Academy Sciences
Shenzhen, China
h.wang@siat.ac.cn

Fangyu Liu*
Shenzhen Institutes of Advanced Technology, Chinese Academy Sciences
Shenzhen, China
University of Chinese Academy of Sciences
Beijing, China
fy.liu1@siat.ac.cn

Xiang Li
School of Electronics and Information Engineering, South China Normal University
Foshan, China
xiang.li2@siat.ac.cn

Huazhen Huang
Shenzhen Institutes of Advanced Technology, Chinese Academy Sciences
Shenzhen, China
hz.huang@siat.ac.cn

Ye Li
Shenzhen Institutes of Advanced Technology, Chinese Academy Sciences
Shenzhen, China
ye.li@siat.ac.cn

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

Abstract

Cross-domain human activity recognition using wearable inertial sensors remains a challenging task, especially in scenarios where no labeled data is available in the target domain. To address this, our team (SIAT-BIT) propose a semi-supervised learning framework based on MixMatch for locomotion and transportation mode recognition. Our approach leverages labeled data from multiple public HAR datasets and unlabeled data from the Sussex-Huawei Locomotion-Transportation Recognition Challenge Task 2 (Kyutech IMU) dataset. The framework integrates pseudo-label generation, data augmentation, soft label sharpening, and cross-sample mixing to mitigate domain shift and label scarcity. Experimental results demonstrate that the proposed method achieves competitive performance in Task 2 with an accuracy of 76.8% and F1 score of 76.5%, confirming its effectiveness in modeling real-world cross-domain activity scenarios without relying on target domain labels.

CCS Concepts

• **Computing methodologies → Semi-supervised learning settings**; • **Human-centered computing → Ubiquitous and mobile computing**.

Keywords

Human Activity Recognition; Semi-Supervised Learning; SHL Dataset; Smartphone

*Both authors contributed equally to this research.
†The corresponding author and supervisor of the team

This work is licensed under a Creative Commons Attribution 4.0 International License.
UbiComp Companion '25, Espoo, Finland
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1477-1/2025/10
https://doi.org/10.1145/3714394.3756209

ACM Reference Format:
Hao Wang, Fangyu Liu, Xiang Li, Huazhen Huang, Ye Li, and Fangmin Sun. 2025. MixMatch-based Semi-supervised Learning Approach for Cross-domain Locomotion and Transportation Mode Recognition. In *Companion of the the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25), October 12–16, 2025, Espoo, Finland*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3714394.3756209

1 Introduction

With the rapid advancement and broad application of wearable technology and mobile sensing devices, Human Activity Recognition (HAR) has emerged as a key enabling technology for smart healthcare, intelligent transportation, and human-computer interaction. HAR methods based on Inertial Measurement Units (IMUs) have demonstrated greater practicality and flexibility compared to traditional vision-based or ambient sensing approaches due to its advantages such as lightweight design, low cost, low power consumption, and independence from external environmental conditions [14]. By modeling and analyzing multi-modal time-series signals collected from accelerometers, gyroscopes and magnetometer, HAR systems can accurately recognize daily activities, modes of transportation, and physical movements. which provides important technical support for applications including remote health monitoring, rehabilitation assessment, fall detection for the elderly, and personalized behavior analysis [7].

In recent years, deep learning-based models such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Transformers have significantly improved HAR performance. However, these methods usually rely on large amounts of labeled data, which are often expensive and time-consuming to collect in real-world scenarios [8]. Therefore, Semi-Supervised Learning (SSL) has received increasing attention, which aim to leverage both labeled and unlabeled data to improve recognition accuracy under limited annotation conditions.

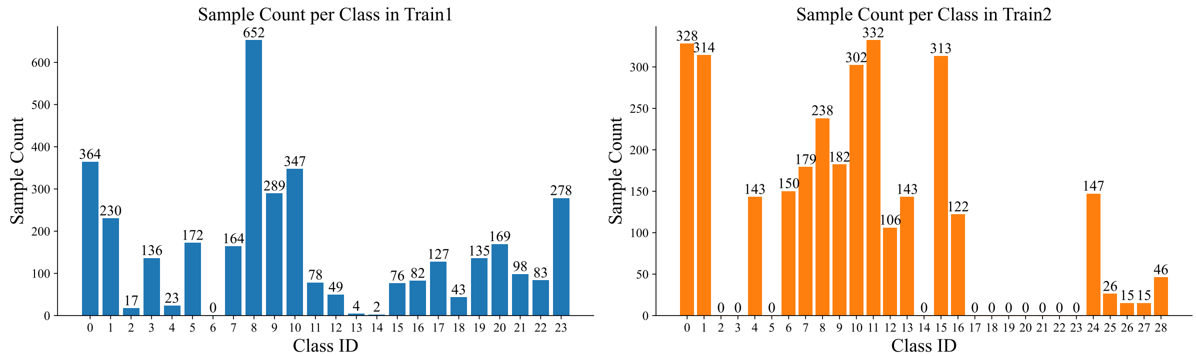


Figure 1: Sample distribution per class in the Train1 and Train2 subsets of SHL Challenge 2025 Task 2.

Table 1: Overview of datasets used.

Locomotion and transportation mode recognition across dataset domains						
Dataset	Year	Subjects	Samples	Classes	Environment	References
PAMAP2	2012	9	17235	18	Semi Free-living	Reiss et al. [11]
UIC HAR	2013	30	10299	6	Lab	Anguita et al. [1]
mHealth	2014	10	2744	12	Lab	Banos et al. [3]
RealWorld	2016	15	71427	8	Semi Free-living	Sztyler et al. [13]
KU-HAR	2021	90	20750	18	Semi Free-living	Sikder et al. [12]
SHL Task1	2025	3	317587	8	Free-living	SHL 2025
KyutechIMU	2025	3	94997	5	Free-living	Asahi et al.[2]
SHL Task2 Test	2025	Unknow	2860	23	Unknown	SHL 2025
SHL Task2 Train1	2025	Unknow	3618	23	Unknown	SHL 2025
SHL Task2 Train2	2025	Unknow	2958	17	Unknown	SHL 2025

Although SSL offers a promising solution to the label scarcity problem, its application in HAR still faces real-world challenges. First, the distributions of data collected under different conditions may vary significantly, leading to domain shifts and degraded model performance. Second, unlabeled data may contain noise, class imbalance, or even unknown activity patterns, which complicates pseudo-labeling. Therefore, designing robust and domain-adaptive SSL methods remains an problem to be solved of HAR.

To advance research in real-world activity recognition, the SHL Recognition Challenge 2025 introduces a Task 2. In this task, participants must develop a recognition model without using any labeled data from the target domain. The target domain consists of the Kyutech IMU data, collected by Kyushu Institute of Technology, which contains large-scale unlabeled inertial signals. Participants are allowed to use publicly available labeled datasets as source domain data to train and initialize the models. This task simulates real-world conditions where models must be adapted to new environments or users without additional annotation.

To address the task, this paper proposes a semi-supervised learning framework based on MixMatch. The proposed approach leverages publicly available labeled datasets as source data and leverages key components of MixMatch including pseudo-label generation, data augmentation, soft label sharpening and cross-sample mixing, effectively mitigates the challenges posed by domain shift and label scarcity. We designed a unified training pipeline that integrates multiple labeled sources with large-scale unlabeled data provided by task2, enabling end-to-end model training without access to target domain labels. Experimental results demonstrate that the proposed framework achieves competitive performance across diverse locomotion and transportation mode recognition tasks, confirming its effectiveness and robustness in complex cross-domain scenarios.

2 Materials and methods

2.1 Datasets

We constructed a diverse training set by integrating several publicly available human activity recognition datasets with varying sensor configurations, subjects, sampling frequencies, and activity categories, including PAMAP2 Dataset [11], UCI HAR Dataset [1], mHealth Dataset [3], RealWorld Dataset [13], KU-HAR Dataset [12], Kyutech IMU Dataset [2], and SHL Dataset [6, 15]. These datasets serve as the labeled source domain for model training, while the Kyutech IMU dataset provided by the organizers is used as the unlabeled target data. The detailed summary of each dataset is shown in Table 1.

Figure 1 shows the sample distribution across categories in the SHL Recognition Challenge 2025 Task 2 training datasets. The left subfigure illustrates the class-wise sample counts in Train1, which covers 24 categories (ID 0–23), while the right subfigure shows Train2, which includes 29 categories (ID 0–28). This difference highlights an inconsistency in label category definitions between the two training datasets. Furthermore, Train1 exhibits severe class imbalance, with some categories (ID 2, 4, 13, 14) having very few samples and category 6 completely missing. In Train2, multiple

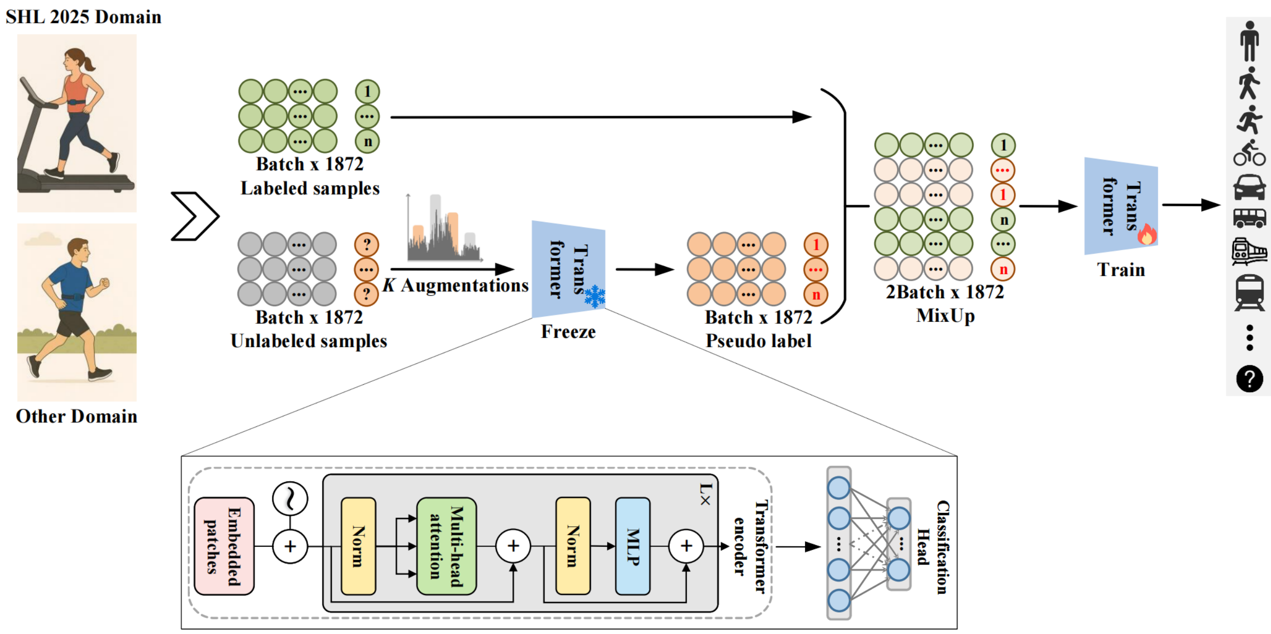


Figure 2: Overview of the proposed recognition pipeline.

Table 2: Overview of computed features per channel.

Domain	Amount	Features
Statistical	31	Absolute energy, Average power, [ECDF Percentile Count_0, ECDF Percentile Count_1], [ECDF Percentile_0, ECDF Percentile_1], [ECDF_0, ECDF_1, ..., ECDF_9], Entropy, Histogram mode, Interquartile range, Kurtosis, Max, Mean, Mean absolute deviation, Median, Median absolute deviation, Min, Peak to peak distance, Root mean square, Skewness, Standard deviation, Variance
Temporal	14	Area under the curve, Autocorrelation, Centroid, Mean absolute diff, Mean diff, Median absolute diff, Median diff, Negative turning points, Neighbourhood peaks, Positive turning points, Signal distance, Slope, Sum absolute diff, Zero crossing rate
Spectral	111	Fundamental frequency, Human range energy, [LPCC_0, LPCC_1, ..., LPCC_9], [MFCC_0, MFCC_1, ..., MFCC_11], Max power spectrum, Maximum frequency, Median frequency, Power bandwidth, Spectral centroid, Spectral decrease, Spectral distance, Spectral entropy, Spectral kurtosis, Spectral positive turning points, Spectral roll-off, Spectral roll-on, Spectral skewness, Spectral slope, Spectral spread, Spectral variation, [Spectrogram mean coefficient_X Hz], [Wavelet absolute mean_X Hz], [Wavelet energy_X Hz], Wavelet entropy, [Wavelet standard deviation_X Hz], [Wavelet variance_X Hz],

Tablenotes: X Hz means at different frequencies. A total of 156 features are extracted from each of the 12 channels, resulting in 1872 features per sample.

categories (ID 2, 3, 5, 14, 17–23) contain zero samples. These inconsistencies and missing classes pose a significant challenge for supervised or semi-supervised learning, requiring careful preprocessing, resampling, and category alignment strategies to ensure model stability and generalization across domains.

2.2 Data preprocessing

We designed a unified data preprocessing pipeline tailored for inertial signal-based activity recognition. All raw IMU signals, including triaxial acceleration, gyroscope, and magnetometer data from multiple datasets and body locations, were first segmented into 500 points windows. Each segmented signal was then processed using the Time Series Feature Extraction Library (TSFEL) [4], a Python package designed for efficient and automatic extraction of time series features.

TSFEL extracts features across statistical, temporal, and spectral domains, totaling 156 features per signal axis (31 statistical, 14 temporal, and 111 spectral features), as shown in Table 2. This results in a 1872-dimensional feature vector per sample for 12-channel input (3-axis accelerometer, gyroscope, magnetometer, and their respective magnitude). This feature extraction scheme was also employed in our previous work [9], demonstrating its effectiveness for time-series classification tasks.

2.3 Model

Existing supervised learning methods can only be applied to data with real labels, which makes the unlabeled data of the Kyutech IMU dataset provided in Task 2 of this competition unusable. This leads to significant limitations in the model's learning of the Kyutech IMU dataset domain. Moreover, the categories of existing labeled datasets are difficult to directly and accurately match with the categories of SHL Recognition Challenge Task 2. Although conventional semi-supervised or weakly supervised methods can use datasets from multiple different domains at the same time, the large differences between different dataset domains will reduce the generalization and robustness of the model.

We constructed a semi-supervised method for cross-domain learning based on MixMatch [5] and Transformer. The data enhancement, soft label sharpening, and cross-sample mixing (MixUp), operations in the MixMatch method enable the model to generalize and learn the mixing distribution of cross-domain samples. These advantages enable our model to simultaneously learn the distribution of unlabeled data and the distribution of data between different dataset domains.

Figure 2 illustrates the overall cross-domain semi-supervised learning framework based on MixMatch and Transformer. It shows the process of pseudo-label generation, data augmentation, soft label sharpening and cross-sample mixing used to train the final classifier.

The overall procedure based on MixMatch is summarized in Algorithm 1. We first apply data augmentation to unlabeled samples by adding Gaussian noise. Pseudo-labels are generated by inference using a frozen model, and further processed by averaging predictions across K augmentations and applying label sharpening. Finally, these pseudo-labels are mixed with samples having ground-truth labels through the MixUp operation to obtain the final pseudo-labels for training. We set the batch size to 64, the sharpening temperature $T = 0.5$, the MixUp parameter $\alpha = 0.75$, and the number of augmentations $K = 2$.

Algorithm 1 A Semi-supervised Learning Approach Based on MixMatch.

Input: A batch of labeled samples $\mathcal{X} = \{x_b, y_b\}_{b=1}^{Batch}, y_b \in \{0,1\}^{Class}(one-hot)$, a batch of unlabeled samples $\mathcal{U} = \{u_b\}_{b=1}^{Batch}$, the sharpening temperature T, the parameter α that controls the Beta distribution, the number K of augmentations.

Output: The prediction results $\tilde{y}$ of mixed samples with labeled samples and unlabeled samples in the current batch, the gradient loss $\mathcal{L}$ for the current batch.

$\bar{p}_b = \frac{1}{K}\sum_{k=1}^{K} softmax(f(Augment(u_b^k)))$

$\hat{p}_b \leftarrow Sharpen(\bar{p}, T)_i = \frac{\bar{p}_i^{\frac{1}{T}}}{\sum_{j=1}^{Class} \bar{p}_i^{\frac{1}{T}}}$ //Pseudo-label

$\bar{\mathcal{X}} = \{x_b, y_b\}_{b=1}^{Batch} \cup \{u_b, \hat{p}_b\}_{b=1}^{Batch}$ //Concat

$\lambda \sim Bata(\alpha, \alpha)$

$\tilde{x}_i = \lambda x_i + (1-\lambda)x_j\ (j \sim Uniform(1, |\bar{\mathcal{X}}|)//MixUp$

$\tilde{y}_i = \lambda y_i + (1-\lambda)y_j\ (j \sim Uniform(1, |\bar{\mathcal{X}}|)//MixUp$

$\mathcal{L}_x = \frac{1}{|\mathcal{X}|}\sum_{x,y\in\mathcal{X}} H(y, f(\tilde{x}))$ //Labeled losses

$\mathcal{L}_u = \frac{1}{|\mathcal{U}|}\sum_{x\in\mathcal{U}} \|\hat{p} - softmax(f(\tilde{x}))\|_2^2$

$\mathcal{L} = \mathcal{L}_x + \lambda_u \mathcal{L}_u$ //Labeled and Unlabeled losses

return $\tilde{y}, \mathcal{L}$

3 Experiment and analysis

3.1 Experimental settings

Since it was uncertain whether the test dataset included samples labeled as class 6, we formulated a 24-class classification task, where the category labels ranged from 0 to 23. We used the usual accelerometer data obtained in the manner ($a_x + gF_x$, $a_y + gF_y$, $a_z + gF_z$) suggested by the organizer, which incorporates gravity-compensated acceleration.

After combining the two training sets of Task 2, we found that the label sample sizes of categories 2, 13, 14, 18, 21, 22 were fewer than 100. In particular, the sample size of label categories 13 and 14 is too small to complete the division of training set, validation set, and test set. To mitigate the empty class problem caused by such data imbalance, we resampled the original data before dividing the data set, and resampled the subcategories with less than 100 samples to make up the number.

After resampling, we divided all 24 subcategories into training set, validation set, and test set according to 8:1:1. As for the data of Task 1, Kyutech IMU dataset, and datasets in other domains that cannot directly and accurately obtain label categories, we regard them as unlabeled data and add them to our semi-supervised algorithm.

Since each batch of our semi-supervised training method based on MixMatch requires the same amount of labeled data and unlabeled data, we recycle the labeled data with less data. This recycling is not a simple duplication: through operations such as data augmentation, label sharpening, and cross-sample mixing, the model dynamically synthesizes new fused examples that reflect the mixed distribution across domains. This improves domain generalization beyond the original sample set.

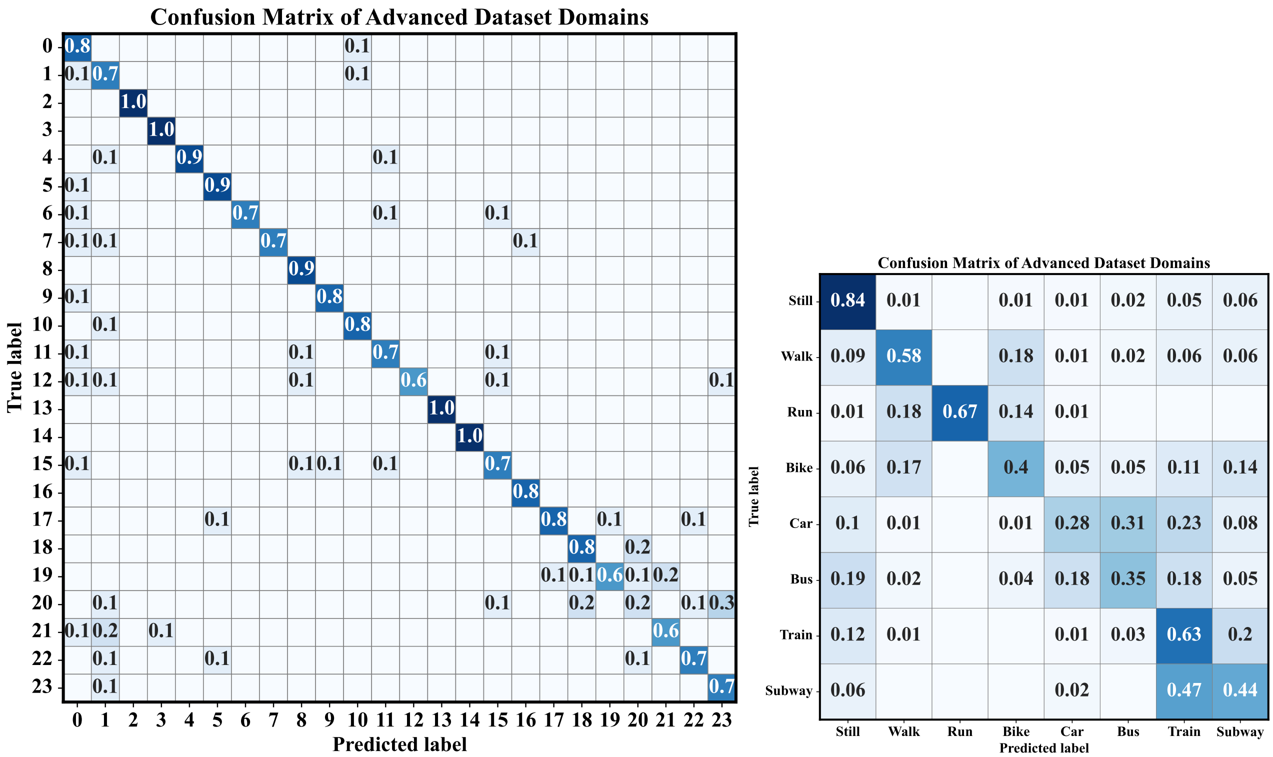
The experiments were conducted on a computing setup with a 12-core 3.7GHz CPU, 128GB RAM, and an NVIDIA RTX 3090 GPU. All implementations were done in Python using Scikit-learn libraries. We used the Adam optimizer for model training, with a fixed learning rate of 0.0001. The batch size was set to 64, and all experiments were trained for 30 epochs. During training, we uses cross-entropy loss for labeled data and predicted results, and the MSE loss for pseudo-labeled and predicted results for unlabeled data.

We use the scikit-learn library to calculate multi-class classification performance metrics with sample weights. The six performance metrics are accuracy (ACC), precision (PRE), recall (REC), F1-Score (F1), Matthews correlation coefficient (MCC), and area under the receiver operating characteristic curve (AUC), where the AUC is derived from a weighted average of the AUCs of all subcategories based on the one-vs-rest algorithm.

3.2 Improving the Test Score

To improve the test performance, we applied several additional strategies during the model retraining and test phase. The adopted enhancements include:

(1) We included the test set of Task 2 in the semi-supervised training process, treating it as an unlabeled dataset. This allowed the model to indirectly adapt to the distribution of the test data without using any labels.
(2) In addition to the SHL Recognition Challenge 2025 dataset and the Kyutech IMU dataset, we try to explored other datasets from different domains for learning cross-domain information. This enriched the domain diversity of the training data and helped the model better handle distribution shifts across unseen domains.



(a) (b)

Figure 3: Confusion matrix: (a) Confusion matrix obtained from Task 2, (b) Confusion matrix obtained from Task 1.

Table 3: Experimental results on Task 2.

Dataset domains	Locomotion and transportation mode recognition across dataset domains							
	Advanced domain samples {Train: Cycle(5344); Val: 668; Test: 668; Unlabeled: 1212731;}							
	ACC	PRE	REC	F1	MCC	AUC	FLOPs	Params
Basic	0.763	0.766	0.763	0.762	0.746	0.938	0.749K	2.992M
Enhanced	0.732	0.736	0.732	0.729	0.713	**0.946**	0.749K	2.992M
Advanced	**0.768**	**0.772**	**0.768**	**0.765**	**0.751**	0.939	0.749K	2.992M

Basic dataset domains: SHL Challenge 2025; Enhanced dataset domains: SHL Challenge 2025 + Kyutech IMU; Advanced dataset domains: SHL Challenge 2025 + Kyutech IMU + Others (i.e. PAMAP2 Dataset, UCI HAR Dataset, mHealth Dataset, RealWorld Dataset, and HAR Dataset); FLOPs is the computational cost per sample inference;

Table 4: Experimental results on Task 1.

Dataset domains	Locomotion and transportation mode recognition across dataset domains							
	Advanced domain samples {Train: 784288; Val: 57576; Test: 57580; Unlabeled: Cycle(319614);}							
	ACC	PRE	REC	F1	MCC	AUC	FLOPs	Params
Basic	0.526	0.541	0.526	0.523	0.441	0.881	0.716K	2.976M
Enhanced	0.546	0.555	0.546	0.544	0.465	0.887	0.716K	2.976M
Advanced	**0.549**	**0.581**	**0.549**	**0.546**	**0.472**	**0.902**	0.716K	2.976M

Basic dataset domains: SHL Challenge 2025; Enhanced dataset domains: SHL Challenge 2025 + Kyutech IMU; Advanced dataset domains: SHL Challenge 2025 + Kyutech IMU + Others (i.e. PAMAP2 Dataset, UCI HAR Dataset, mHealth Dataset, RealWorld Dataset, and HAR Dataset); FLOPs is the computational cost per sample inference;

3.3 Classification performance

As can be seen from Table 3, the addition of the Kyutech IMU dataset causes a reduction in model accuracy on the test set derived from the SHL Challenge 2025 Task 2. This is due to the domain discrepancies introduced by the Kyutech IMU dataset, which weakens the model's ability to fit the distribution of Task 2. In contrast, when more domain datasets are added, the model's performance improves again. This is because the inclusion of multiple domains allows the model to learn from more diverse data distributions, thereby enhancing its generalization and robustness.

In Figure 3(a), we observe that categories 2, 13 and 14 exhibit high accuracy because their originally small sample size, which was expanded through resampling. However, the small sample size may also leads to a large gap between the prediction results obtained from these categories and the actual prediction performance. According to the confusion matrix, the category 20 has the worst prediction effect. Since category 20 has a high similarity to category 18 and category 23, many samples with real label 20 are mispredicted as categories 18 or 23.

Since the Task 2 of SHL Recognition Challenge 2025 contains relatively few labeled samples, the prediction results may not fully reflect the actual prediction performance. We applid our method to Task 1 to more realistically reflect the prediction performance of the model and the role of multiple dataset domains. We use the training data of the four parts of Task 1 as labeled training samples, and split the validation data of the four parts into labeled validation samples and test samples at a 1:1 ratio according to the number of subcategories. The test data of Task 1, data from Task 2, and other domin dataset were regarded as unlabeled samples.

As shown in Table 4, it can be found that the model's performance slightly improves with the addition of datasets from different domains. This shows that the model has learned the distribution of data from different dataset domains, which contributes to better generalization and robustness. Figure 2(b) shows that in Task 1, the model performs well on categories such as Still, Run, and Train, but underperforms on categories like Bike, Car, and Bus.

Despite leveraging a semi-supervised cross-domain learning framework, the proposed model maintains high computational efficiency. As shown in Tables 3, the model contains only 2.99M parameters and requires approximately 0.75K FLOPs per inference. This lightweight architecture makes it highly suitable for real-time deployment on edge devices with limited computing resources.

4 Conclusion

In this paper, we present a MixMatch-based semi-supervised learning approach for cross-domain locomotion and transportation mode recognition. By leveraging multiple labeled source datasets and large-scale unlabeled target data from Kyutech IMU, our method addresses the challenge of domain shift and the lack of labeled data in SHL Recognition Challenge 2025 Task 2. The unified training strategy combines key elements of semi-supervised learning including pseudo-label generation, data augmentation, soft label sharpeningand cross-sample mixing, significantly improving generalization and robustness of the model. Experimental results show that our method achieves an accuracy of 76.8% on Task 2 when integrating additional domain datasets, demonstrate that our approach not only adapts well to target-domain distribution without requiring labels, and shows improved generalization when incorporating additional domain data. The recognition result for the testing dataset will be presented in the summary paper of the challenge [10].

5 Acknowledgments

This work was supported in part by the Key Research and Development Plan of Guangdong Province under Grant 2022B1515120062; in part by the Shenzhen International Cooperation Project under Grant GJHZ20220913142808016; in part by the Shenzhen Sustainable Development Special Project under Grant KCXFZ202307310941-00001 and KCXFZ20240903094300001; in part by the Joint Fund of NSFC and Chongqing under Grant U21A20447; and in part by the Joint Fund of Yeqisun and NSFC under Grant U2241210; in part by Guangzhou Key R&D Program under Grant 202206010127 and 2023B03J1341, and in part by National Natural Science Foundation of China under Grant 62401215.

References

[1] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, Jorge Luis Reyes-Ortiz, et al. 2013. A public domain dataset for human activity recognition using smartphones. In *Esann*. 3–4.

[2] Miyazaki Asahi, Tengjiu Huang, Okita Tsuyoshi, and Nishikawa Asahi. 2025. Acquisition of Unlabeled Dataset for Human Activity Recognition. *IPSJ(UBI)* (2025).

[3] Oresti Banos, Rafael Garcia, Juan A Holgado-Terriza, Miguel Damas, Hector Pomares, Ignacio Rojas, Alejandro Saez, and Claudia Villalonga. 2014. mHealthDroid: A novel framework for agile development of mobile health applications. *Ambient Assisted Living and Daily Activities* 8868, 14 (2014), 91–98.

[4] Marília Barandas, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, and Hugo Gamboa. 2020. TSFEL: Time series feature extraction library. *SoftwareX* 11 (2020), 100456.

[5] David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A Raffel. 2019. Mixmatch: A holistic approach to semi-supervised learning. *Advances in neural information processing systems* 32 (2019).

[6] Hristijan Gjoreski, Mathias Ciliberto, Lin Wang, Francisco Javier Ordonez Morales, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2018. The university of sussex-huawei locomotion and transportation dataset for multimodal analytics with mobile devices. *IEEE Access* 6 (2018), 42592–42604.

[7] Pranjal Kumar, Siddhartha Chauhan, and Lalit Kumar Awasthi. 2024. Human activity recognition (har) using deep learning: Review, methodologies, progress and future research directions. *Archives of Computational Methods in Engineering* 31, 1 (2024), 179–219.

[8] Jiaxi Li, Zhelong Wang, Zheng Wang, Sen Qiu, Daoyong Peng, Ke Zhang, and Fang Lin. 2024. A comprehensive evaluation method for frailty based on semi-supervised learning and transfer-learning. *Information Fusion* 111 (2024), 102504.

[9] Fangyu Liu, Hao Wang, Weilin Zang, Ye Li, and Fangmin Sun. 2024. Influencing Factors Mining and Modeling of Energy Expenditure in Running Based on Wearable Sensors. In *Proceedings of the 2024 International Conference on Sports Technology and Performance Analysis*. 38–46.

[10] Tsuyoshi Okita, Ukita Kosuke, and Miyazaki Asahi. 2025. Foundation Models to Tackle Activity Recognition in Unknown Domain: Sussex-Huawei Locomotion Challenge 2025 Task 2. In *Proceedings of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2025 ACM International Symposium on Wearable Computers*.

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

[12] Niloy Sikder and Abdullah-Al Nahid. 2021. KU-HAR: An open dataset for heterogeneous human activity recognition. *Pattern Recognition Letters* 146 (2021), 46–54.

[13] Timo Sztyler and Heiner Stuckenschmidt. 2016. On-body localization of wearable devices: An investigation of position-aware activity recognition. In *2016 IEEE international conference on pervasive computing and communications (PerCom)*. IEEE, 1–9.

[14] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. 2019. Deep learning for sensor-based activity recognition: A survey. *Pattern recognition letters* 119 (2019), 3–11.

[15] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2019. Enabling reproducible research in sensor-based transportation mode recognition with the Sussex-Huawei dataset. *IEEE Access* 7 (2019), 10870–10891.