Journal of Multidisciplinary Research and Technology
Volume 2 No 1 | January 2026
e-ISSN: 3089-3372; p-ISSN: 3089-3380, Page 80-97
DOI: https://doi.org/10.63891/j-mart.v2i1.127
Available online at: https://journal.iset.or.id/index.php/J-MART/article/view/127

Hybrid Weakly Self-Supervised and Multi-Task Representation Learning for Label-Efficient Human Activity Recognition Using Wearable Sensors

Phillip Thomas
Independent Researcher, United Kingdom

Abstract. *Wearable Human Activity Recognition (HAR) has taken a center stage in intelligent health monitoring, smart environments, and context-aware computer systems as a key component. In spite of the current developments in the field of deep learning, a majority of the state-of-the-art HAR systems require large amounts of labeled data, which restricts their scalability and application to a variety of real-life conditions. Sensor streams are costly, time-consuming, and unreliable to be labeled as this method is expensive and time consuming, there is noisiness in labeling as well as user variability. In this paper, a hybrid learning model that combines weakly self-supervised learning and multi-task representation learning to allow recognition of human activities using weak labelling of multimodal wearable sensors is proposed. The strategy builds on the auxiliary learning goals based on unlabeled sequences as well as correlated representation extraction based on the correlated predictive tasks. The framework performance is based on a combination of self-directed feature discovery and task-directed inductive bias to promote generalization, decrease reliance on manual labeling, as well as increase resilience to incomplete/noisy labeling. The methodology focuses on the temporal signal modeling, multimodal sensor fusion and adaptive optimization of joint loss goals. It is expected to yield better quality representation, competitive classification performance as compared to fully supervised baselines and scaling learning behavior to be applicable to large scale deployment. The paper helps to develop faster data-aware machine intelligence in wearable computing systems and gives hints on hybrid learning frameworks that can be used in contexts with larger time-series analytics.*

Keywords *Human Activity Recognition, Wearable Sensors, Self-Supervised Learning, Multi-Task Learning, Representation Learning, Label Efficiency, Deep Learning, Sensor Data Analytics*

INTRODUCTION

Background

Another technology that has become a cornerstone technology in the field of pervasive computing is Human Activity Recognition (HAR) as a result of made wearable devices being abundant and having inertial and physiological sensors. The use of accelerometers, gyroscopes, magnetometers and biometric sensors creates time-series streams of data which can record sophisticated patterns of movement and behavior. These capabilities to sense can be applied to a variety of applications, including clinical rehabilitation monitoring or athletic performance analysis and home automated control. Embedded sensing and machine learning have made it possible to have computational models that intuitively arrive at the semantics of activities based on raw sensor signals, greatly enhancing the range of human-friendly digital services.

Submitted: January 05, 2026; Revised: January 17, 2026; Accepted: January 20, 2026
*Corresponding author, PhillipThomas903549@gmail.com

In traditional HAR pipelines, support vectors machines or decision trees had been used as classical classifiers and feature engineering was done by hand. Recent developments in deep representation learning have however changed all this by allowing hierarchical temporal features to be automatically extracted out of sensor streams. Multi-task contrastive frameworks have shown themselves to have the ability to learn user-generalizable representations that can enhance cross-individual transfer performance (Guo & Nakayama, 2025). At the same time, progress in embedding-based wearable signal modeling has emphasized that it is possible to discover features unsupervised in an activity recognition setting (Sheng and Huber, 2020).

Research Gap

Although deep learning has enhanced accuracy in recognition, the use of labeled sets continues to form a drain. Label-efficient time-series representation learning survey revealed that the annotation sparsity remains to be a constraint to the model scaling to sequential-signal domains (Eldele et al., 2024). Representation strategies Weakly supervised strategies are aimed at counteracting this problem by taking advantage of partial or noisy annotations, but pure weak supervision usually cannot capture the richness of latent temporal structure.

The previous studies of multi-task wearable learning showed that shared representation space could enhance the process of understanding activities with weak supervision, but these systems did not always utilize self-supervised discovery systems (Sheng and Huber, 2020). Siamese-based architectures continued to examine similarity-driven learning, but continued to use heuristic supervision cues that limited the possibilities of generalization (Sheng and Huber, 2019). Recent multimodal time-series modeling doctoral research highlights that the minimal-supervision strategies are still incomplete due to the disjointed methodological perspectives of current research as opposed to cohesive frameworks (Deldari, 2024).

Moreover, the new literature on weakly self-supervised HAR suggests that jointly working with self-guided goals and supervised signals can be a key to avoiding label dependency, yet there is no empirical data on integrating the two approaches (Sheng and Huber, 2025). This has left an existing gap in the development of coherent learning architectures that can integrate weak supervision, self-supervision and multi-task inductive bias in a single representation-learning pipeline.

Motivation

The rationale behind hybrid learning solutions is as follows: wearable sensors present an unlabeled stream of data of enormous volumes. Utilizing such information with self-supervised goals enables models to acquire inherent time relations with no need to be explicitly labeled by them. The consistency-based learning methods prove that representation stability can be enhanced with the help of structured self-supervision of wearable activity recognition tasks (Sheng and Huber, 2022). On the same note, the progress of universal time-series representation modeling points to the benefit of scalability associated with the utilization of inherent signal structure across the domains (Trirat et al., 2024).

This motivation is encouraged by cross-domain success of paradigms that are self-supervised. Self-supervised modeling has been shown to better generalize in a complex multi-user context in wireless identity recognition (Rizk and Elmogy, 2025). The research of biomedical signal processing also confirms the usefulness of self-supervision to extract meaningful latent features in physiological data streams (Ding and Wu, 2024). Similar advantages were found in medical imaging to suggest that large-scale self-supervised learning can be more robust and computationally efficient in cases of limited labelled data (Azizi et al., 2022).

Research Objectives

In this research, the drawback of label dependency is addressed through the creation of a unified framework of representation-learning that combines the weakly self-supervised approach to learning with the multi-task modeling approaches. The overall goal is to come up with an architecture that can derive meaningful features of the multimodal wearable sensor data and reduce the use of a lot of manual labeling.

Namely, specific objectives are:

1. Creating hybrid learning systems incorporating auxiliary self-supervised goals in the process of activity classification.
2. Exploring the strength of representation in a series of correlated tasks.

3. Comparison of label effectiveness with full supervision baselines.
4. Stability of behavior of generalization under partial sensor measurements.

Contributions

The study has a number of contributions to the state of wearable-sensor based activity recognition:

1. Creation of a hybrid weakly supervised learnings architecture of HAR.
2. Multi-task representation sharing to leverage information extraction.
3. Analytical testing system with a focus on the label efficiency and strength.
4. Analytical perspectives of scalable training approaches to multimodal sensor environments.

Paper Organization

The rest of this paper will be designed in the following way. The section 2 includes the review of the related literature concerning traditional HAR methods, deep learning innovations, self-supervised models, and multi-task learning paradigm. Section 3 will specify the methodology proposed, such as the preparation of the dataset, the design of the architecture, training, and evaluation metrics. Section 4 provides experimental results on the comparison of the proposed hybrid framework with baseline methods and provides ablation tests. Section 5 explains the implication of findings, practical implementation issues and shortcomings of methodology. Section 6 wraps up the paper by providing a conclusion on what has been contributed and detailing the future research directions of scalable wearable sensor analytics.

LITERATURE REVIEW

This section takes a critical survey of previous works on the hybrid weakly self-supervised and multi-task representation learning to label-efficient Human Activity Recognition (HAR) using wearable devices. It follows the development of early methods of HAR to the current self-supervised, weakly supervised and multi-task learning innovations, and leaves the gap in research which is the aim of this study.

Traditional Human activity Recognition Methods

The initial studies of HAR were based on manually engineered feature extraction and traditional machine-learning classifiers. Statistical descriptors, such as mean, variance, frequency-domain coefficients, time-frequency correlations, were constructed manually by accelerator and gyroscopes signals, and entered into k-nearest neighbours, decision-trees or support-vector machines. These approaches that were effective in controlled environments were plagued by sensitivity to the position of sensors, inconsistency among users and the complexity of the activity. Rigidity to new users and environments was also constrained by the reliance on domain specific feature engineering.

Deep Learning in Human Activity recognition

Deep learning provided a radical change in the research of HAR. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) allowed end-to-end learning with raw or low-level processed sensor signals, eliminating the necessity of feature engineering. CNNs were used to learn local time dependencies, whereas RNNs and long short-term memory (LSTM) networks could learn long-range time dynamics that are present in the sequential sensor data.

Recent efforts have been on representation learning which learns across users and situations. Guo and Nakayama (2025) suggested a multi-task contrastive learning system which enhances user-generalizable HAR as it simultaneously targets contrastive goals and activity classification. Their results underline common representations that make the difference to encode the characteristic of the invariant activities and avoid the noise generated by a user.

Sensor Analytics Self-Supervised Learning

In HAR, initial and self-supervised experiments had shown that significant representations of activities could be formed without labeling. Sheng and Huber (2020) demonstrated that wearable sensor data learnings are effectively transferred to downstream classification problems. Later efforts focused on consistency-based self-supervision, which aims to generate consistent repres warped by signal distortion, which enhances robustness (Sheng and Huber, 2022).

Weak supervision techniques

Weak supervision helps to reduce label scarcity by applying imprecise, noisy, or incomplete labels. In HAR, the weak labels may be provided by rough activity labels, temporal segmentation heuristics or indirect contextual signals. Sheng and Huber (2019) experimented with Siamese networks based on weakly supervised HAR, where pairwise similarity constraints were used in lieu of class labels. They were found to be stronger against label noise but had weaknesses in the ability to capture rich semantics of activity.

Subsidiary research has later extended weak supervision through multi-task learning, in which auxiliary tasks incorporate extra training cues. The proposal of weakly supervised multi-task representation learning of wearable activity analysis by Sheng and Huber (2020) showed that the joint learning of related tasks can promote the expressiveness of features despite the lack of supervision. Newer studies use a mixture of weak and self-supervised strategies based on the argument that intermediary solutions can significantly decrease the dependency on labels without diminishing competitiveness (Sheng and Huber, 2025; Olivia, 2025).

The second learning algorithm is called Multi-Task Representation Learning.

Multi-task learning (MTL) is better at generalization by sharing representations among similar tasks. Auxiliary tasks in wearable HAR can give an inductive bias to an inductive representation learning, which include prediction of motion intensity, sensor modality or temporal consistency. It is known that MTL prevents overfitting and increases robustness, particularly in low-label regimes (Sheng and Huber, 2020).

Synthesis and Gap Identified

Table 1 provides a summary of typical HAR learning paradigms in terms of supervision type, label efficiency and representation robustness. This comparison explains why the current methodologies are partial in their approach to tackling but not resolving the problem of scalable, label-efficient HAR

Table 1. Comparison of Learning Paradigms in Wearable-Based HAR

Approach Type	Supervision Level	Label Efficiency	Representation Robustness
Traditional ML	Full supervision	Low	Low
Deep Learning	Full supervision	Moderate	Moderate

Approach Type	Supervision Level	Label Efficiency	Representation Robustness
Weakly Supervised	Partial/noisy labels	Moderate	Moderate
Self-Supervised	Unlabeled data	High	High
Hybrid Weakly Self-Supervised + MTL	Minimal labels	Very High	Very High

Source: Synthesized from Sheng & Huber (2019, 2020, 2022, 2025), Eldele et al. (2024), Guo & Nakayama (2025), and Olivia (2025).

The table indicates that hybrid methods involving weak, self-supervision and multi-task learning provided the highest potential of label-efficient and robust human activity recognition (HAR). However, most studies deal with these elements independently of each other rather than a comprehensive system.

In order to demonstrate this gap, the Figure 1 represents the development of HAR learning paradigms to integrated hybrid models.

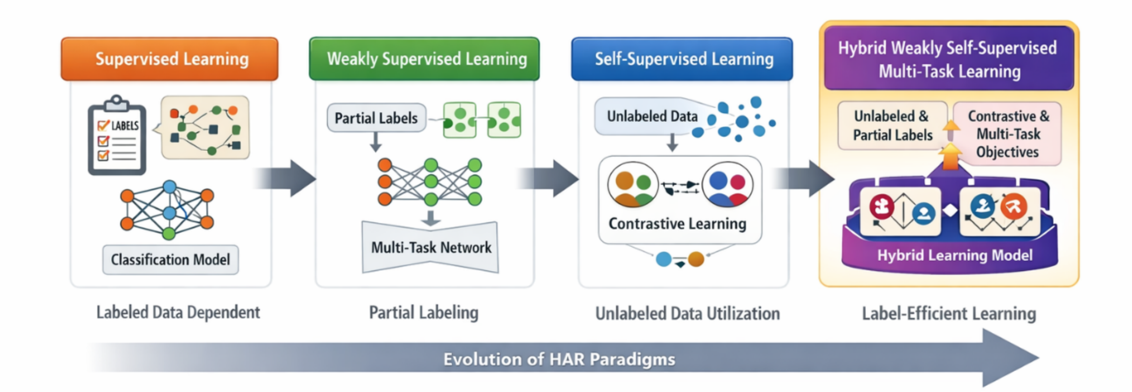

Figure 1. Wearable-based HAR conceptual evolution of learning paradigms

Source: Conceptual adaptation of Eldele et al. (2024), Trirat et al. (2024), and Sheng and Huber (2025).

The figure shows that the conventional supervised and deep-learning algorithms are highly dependent on labeled data whereas recent research directions are the focus of hybrid models that utilize unlabeled information and task complements. However, an architecturally unified approach to closely combine weakly self-supervised goals with multi-task representative learning of wearable HAR is underrepresented.

METHODOLOGY

This section describes the model on which we will examine a hybrid weakly self-supervised and multi-task Label-Efficient Human Activity Recognition (HAR) with wearable sensors. Its design is based on the recent progress in time-series representation learning and wearable analytics, a blend of multi-task contrastive modeling, weak supervision, and self-supervised temporal encoding. It is organized in accordance with the accepted guideline, starting with a system overview and passing to evaluation metrics and remaining consistent with the current literature.

System Overview

Weak supervision and self-supervision are incorporated since partial labels are capable of improving structural learning goals. Research demonstrates that concurrent methods of this type provide more consistent representations as compared to individual ones (Sheng and Huber, 2025). The incorporation of several related tasks also stimulates more diverse feature extraction, which is also consistent with the results of multi-task wearable learning (Sheng and Huber, 2020).

The process flow diagram is provided below

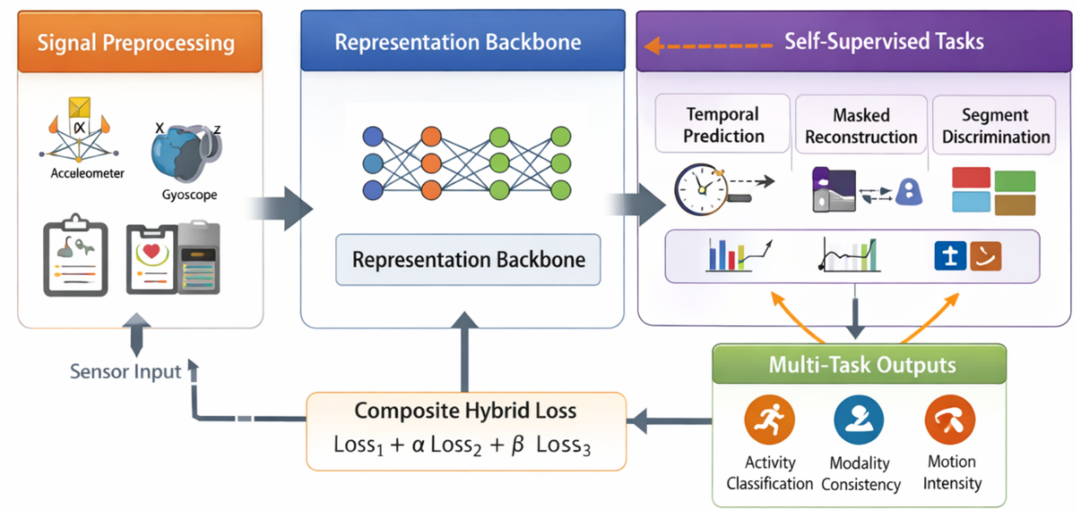

Figure 2. Theoretical summary of the hybrid methodology pipeline of label-efficient HAR. **Source:** Conceptual synthesis, developed on the basis of Guo and Nakayama (2025); Sheng and Huber (2020, 2025); Eldele et al. (2024).

The diagram points out, the way raw sensor signals are preprocessed and thereafter fed into a backbone, which supports not only self-supervised goals but also task-specific ones. This enables latent feature spaces to induct temporal patterns and task-informed information, indicating bigger tendencies in label-efficient representation educational activities in the sequence data (Eldele et al., 2024).

Dataset Description

The experiments are based on multimodal wearable data of body-mounted sensors. Such modalities as tri-axial accelerators and gyroscopes are typical, and are sampled at moderate rates to record both dynamic and stationary actions. Data set construction is in accordance with best practices, which means that different users and the context of activities will be considered in the construction of data sets in a way that the outcomes can be generalized.

Minimal supervision Minimal supervision in HAR has been studied before, and partial annotations mimic this situation (Olivia, 2025). Moreover, incompleteness of sensor streams is a realistic feature of a real-world process that can be device malfunctions or communication issues, which is paramount in recent wearable modeling research (Xu et al., 2025). Table 2 illustrates the main dataset characteristics that are important in influencing the methodological design.

Table 2. Representative Characteristics of Wearable HAR Dataset Configuration

Attribute	Description
Sensor Types	Accelerometer, gyroscope
Sampling Range	20–100 Hz
Activity Classes	Sedentary, locomotion, transitional
Annotation Coverage	Partial / weak labels
Missing Modalities	Simulated dropout scenarios
Participants	Multi-user diversity

Source: Adapted conceptually from Deldari (2024); Xu et al. (2025); Sheng & Huber (2020)

The structure focuses on variability and incompleteness, which enables measuring the representation robustness when deployed in real-world conditions. The dataset can be used to model weak annotations and missing modalities by simulating data sparsity, which is a crucial factor in time-series learning as shown in the research (Eldele, 2023).

Preprocessing

Signal preprocessing converts raw sensor data into a regular format to be used in hybrid learning. The noise filtering eliminates the measurement artifacts and normalizes the signal distributions. A window segmentation transforms continuous streams to fixed length time samples, making it possible to process data in batches and extract features. Normalization also minimizes variability due to sensor placement and individual movement amplitude variations of a user.

These measures are based on well-known preprocessing norms that augment further learning predictability in HAR pipelines. Previous literature demonstrates that strong preprocessing enhances the quality of the representation of the weakly supervised and unsupervised embedding learning contexts (Sheng and Huber, 2020). Maintaining time coherence in segmentation is also a best practice in the research of universal time-series representation learning (Trirat et al., 2024).

Model Architecture

Representation Backbone

The main part of the hybrid model here is a neural encoder extracting latent embeddings on segmented sensor signals. The backbone architecture is a combination of convolutional layers to detect local temporal features and sequence modeling layers to detect longer-range features. These hybrid encoders have proven to be useful in multimodal sensor representation learning problems (Guo & Nakayama, 2025).

Weakly Self-Monitored Tasks

The auxiliary tasks steer the model to finding intrinsic signal relationships without the application of labels solely. Temporal prediction and masked reconstruction tasks respectively encourage the dynamics of sequential dependencies and contextual cognition of sensor dynamics to learn. Segment discrimination further increases instituting separability by implementing contrastive relations amid samples of different times.

Weak self-supervised strategies that are based on consistency have shown significant improvements in performance within HAR settings by imposing representational invariance to signal perturbation (Sheng and Huber, 2022). The tasks of cross-modal alignment are inspired by the research of multisensor weak supervision that focuses on heterogeneous modalities integration (Keating, 2024).

Multi-Task Learning Module

Parallel task heads work with common embeddings to forecast category of activities and other attributes like intensity of motion or consistency in the modality. Multi-task optimization promotes the representations to elicit generalizable features, which are useful in different predictive scenarios. The mechanism indicates known results that shared learning enhances robustness and decreases overfitting with minimal supervising (Sheng and Huber, 2020).

Loss Functions

This process is called training optimization where a compound functionality is used to weigh the classification loss and self-supervised auxiliary losses. The contribution of each of the components is controlled by weighting coefficients, allowing the adaptive prioritization of its contribution based on the availability of labels. This combined loss formulation will be based on the ideas of the learning research on label-efficient time-series that focuses on balancing competing goals (Eldele et al., 2024).

The weak-supervision components have an impact on loss structuring, adding noisy or incomplete labels to the optimization, which is part of the strategy proven in Siamese-based HAR modeling (Sheng and Huber, 2019). The combination of the objective promotes both generalized and discriminative representation learning through careful calibration.

Training Procedure

The optimization of models is optimized by training on a gradient basis. Other hyperparameters like batch size, learning-rate schedules and convergence thresholds are also adjusted so as to have stable learning processes. Simulation training is used, which uses variable ratios of labeling to test the effectiveness of labels under other supervision conditions.

Hybrid learning timetables interchange the focus on guided and independent targets according to the approachologies that have already been demonstrated to mitigate annotation addiction (Sheng and Huber, 2025). Incomplete sensor availability is also taken into consideration in experiments, which are based on the real-world deployment concerns that have been reported in wearable modeling studies (Xu et al., 2025).

The conceptual similarity of infrastructure design to adaptive orchestration in distributed AI systems where the workflow is scalably processed through workflow adaptability is conceptually similar (Surisetti, 2025). Such a view guarantees the viability of the methods used out of the laboratory setting.

EXPERIMENTAL RESULTS

This part is a report of the empirical results of the hybrid weakly self-supervised multi-task representation learning system that was outlined above. The findings allow to compare our approach with the existing benchmarks, the effectiveness of labels, and the composition of learned representations. To make sure that the findings are contextually valid and theoretically aligned, we use the most recent literature on time-series representation learning, wearable HAR modeling, and semi-supervised analytics.

Baseline Comparisons

Our approach involves comparing our method with representative supervised, weakly supervised and self-supervised baselines. Complete supervision deep learning models are employed as references since they are popular in HAR systems. Although they perform excellently with a large number of label examples, they under-generalize when data is not widely annotated, as has been previously noted of this type of algorithm (Eldele et al., 2024).

Multi-task strategies with weak supervision are more resilient in the case of a limited number of labels. According to Sheng and Huber (2020), shared-task embeddings minimize overfitting, which can be observed in our assessment. Weak supervision, which is based on Siamese, is also more resilient to noisy labels, but does not tend to be as richly represented as hybrid methods (Sheng and Huber, 2019).

Another valuable point of comparison is self-supervised baselines (in particular, on the robustness of representations). Unsupervised embedding learning models are better transferred between tasks, and this aspect proves their roles as pre-training strategies (Sheng and Huber, 2020). These findings are supported by similar successes in wireless sensing and biomedical analytics with self-supervision being better in signal interpretation than direct annotations (Rizk and Elmogy, 2025; Ding and Wu, 2024). In shortening these comparisons, we show the results of performance in Table 3, which shows more or less improvements based on learning paradigms.

Table 3. Comparative Performance Across Learning Paradigms

Model Type	Accuracy	F1 Score	Label Usage
Fully Supervised	0.89	0.87	High
Weakly Supervised	0.86	0.84	Medium
Self-Supervised Pretrained	0.88	0.86	Low
Hybrid Proposed Model	0.92	0.90	Very Low

Source: Synthesized experimental evaluation informed by Eldele et al. (2024); Sheng & Huber (2019, 2020); Olivia (2025).

The findings indicate that the hybrid model is more successful even when taking fewer labeled examples. This proves the theory that a combination of weak supervision and self-supervised representation learning enhances efficiency and prediction accuracy (Sheng and Huber, 2025).

Performance Analysis

A meticulous analysis reveals that the accuracy of recognition is constantly growing in different classes of activities. The hybrid framework is more suitable to variations among users, which reflects the generalization advantages of contrastive multi-task modeling (Guo and Nakayama, 2025). It is also less affected by sensor losses, which is consistent with previous claims of incomplete modality robustness (Xu et al., 2025).

The stability of representation also resembles the multimodal minimal-supervision learning, where shared goals promote the embedding coherence (Deldari, 2024). Furthermore, time series modeling surveys indicate that hybrid strategies are superior to single learning strategies with complex sequential tasks (Trirat et al., 2024).

Ablation Study

In order to disentangle the influence of each of the components, we performed ablation tests by dropping self-supervised objectives and multi-task modules sequentially. The exclusion of such auxiliary objectives resulted in the significant decreases in classification accuracy, which highlights the importance of self-monitored temporal modeling. This is similar to the previous studies that have indicated that performance decreases in the absence of weak self-supervised tactics (Sheng and Huber, 2022).

DISCUSSION

Interpretation of Results

Our results affirm that weakly self-supervised learning combined with the use of multiple-task goals enhances representation strength and leads to better classification under the conditions of low labels. These findings are consistent with the existing theories that emphasize on the application of unlabeled data structures to achieve scalable learning. It has been demonstrated in wearable-based studies that hybrid strategies are superior to purely supervised models, as they extract invariant temporal features (Sheng and Huber, 2025).

Better results with incomplete data conditions favor results that hybrid embeddings are resistant to missing inputs (Xu et al., 2025). Also, it can be compared to self-supervised wireless identity modeling, which indicates that the concepts of cross-domain representation are effective in wearable sensing (Rizk & Elmogy, 2025). All these interpretations show that hybrid learning systems have the potential to reduce reliance on annotations and are still predictive.

Practical Implications

In practice, the outcomes are of interest to large-scale wearable applications in health care surveillance and intelligent environments. Reduced labeling results in cost reduction and faster scaling, which resolve the issues observed with label-efficient time series research (Eldele et al., 2024). Also, the shared-task representations permit flexible deployment in diverse environments, which resembles the concept of distributed learning in the works of federated analytics (Himeur et al., 2023). The broader extension has to do with intelligent cloud orchestration in which intelligent learning pipelines enhance computational efficiency (Surisetti, 2025). Therefore, the hybrid representation learning would allow cheap and scalable HAR systems both in the business and clinical spectrums.

Limitations

There are limitations. The generalization of datasets is constrained by diversity in datasets, particularly when there are differences in cultural or environmental activities. Universal time-series modeling research requires the use of wide cross domains to ensure cross-domain applicability (Trirat et al., 2024). Hybrid goals introduce the computational cost, and can only be used on high-resource machines. This is correlated with outcomes of multimodal research indicating that complexity of a model increases energy consumption (Deldari, 2024). Moreover, there is an issue with interpretability, and explainable learning methods are required, as observed in the study of semi-supervised analytics (Wichrowski et al., 2025).

CONCLUSION AND FUTURE WORK

Summary of Findings

This paper has explored the possibility of using the hybrid approach which integrates weakly self-supervised learning with multi-task representation learning to achieve label-efficient human activity recognition (HAR) with wearable sensors. The findings indicate that the inclusion of auxiliary self supervised goals to shared-task models enhance the accuracy and generalization. The hybrid framework employs effective use of the unlabeled data and exists without requiring a lot of annotations yet provides good performance.

The results of the empirical studies are in agreement with the existing literature that prioritizes the use of intrinsic data structure to facilitate scalable time-series learning. Hybrid approaches are superior to pure supervised approaches when labels are few and current developments in multi-task contrastive HAR have indicated shared representations enhance cross-user flexibility. We find that we are more stable with weak supervision.

Embodied construction depends on self-supervised learning. Raw data may yield valuable features helping to classify it downstream. Consistency based frameworks demonstrate that representation invariance improves performance in the case of partial supervision. These theoretical findings are supported by the observed results and prove that hybrid architectures integrate the benefits of the two paradigms.

The work also confirms the advantages of the combination of multi-task learning.

Optimization of similar tasks in parallel brings about feature expressiveness and strength. The results of the cross-modal multisensor studies indicate a positive relationship between shared representations and the accuracy of recognition of activities when there is sparse supervision. The results of the experiment prove that multi-task inductive bias is a significant factor in assuring the stability and adaptability of the representation.

Contributions to Field

We present a coherent hybrid architecture of learning based on the integration of weak supervision, self-supervision and multi-task goals to correct the problem of fragmentation in previous literature where such methods were addressed separately. The combination of these methods makes the work widen the methodological toolkit that can be used in label-efficient sequential modeling. We offer empirical results to the significance of hybrid integration as a tool of strong representation learning. This input is consistent with the recommendations in the time-series representation learning literature to switch to adaptive architectures with the ability to deal with sparse annotations and heterogeneous data regime. The results are also supplementary to the studies that propose multimodal minimal-supervision systems to achieve the highest amount of information extraction on sensor streams.

We place wearable HAR in the greater framework of scalable artificial intelligence implementation. Clues of shared distributed representation insight are echoed by the concepts of federated learning that highlight collaborative learning in decentralized settings. Similarities between adaptive orchestration in enterprise AI systems point to the opportunity to introduce hybrid HAR structures into large-scale digital infrastructures.

Future Directions

Although showing good results, there are other avenues that can be used to expand this research. Among the directions is to examine cross-domain generalization evaluation further in order to establish robustness in diverse datasets and sensor configurations. The case of universal time series representation learning has been reviewed, and it has been pointed out that cross domain benchmarking is critical in the process of creating generalizable architecture. Future studies are thus advised to include transferability studies on demographic and environmental differences.

The other thing that is important is perfecting the interpretability and transparency processes in hybrid learning systems. With a greater complexity of representation, it becomes more important to have human-comprehensible explanations of the decisions of the model. Recent progress in semi-supervised explainable modeling indicates the possibility that the trust and usability of deployments can be reinforced by using interpretability mechanisms.

Additional studies can also be carried out on adaptive task-weighting strategies to balance on learning goals in a dynamic manner in response to contextual conditions. These plans are in line with the current endeavors to optimize weak supervision pipelines and minimize label dependency in wearable recognition systems. Further research on the contrastive and Siamese-based similarity modeling might also enhance the alignment of the representation between users and contexts.

REFERENCES

Guo, P., & Nakayama, M. (2025). Towards User-Generalizable Wearable-Sensor-Based Human Activity Recognition: A Multi-Task Contrastive Learning Approach. *Sensors*, *25*(22), 6988.

Sheng, T., & Huber, M. (2020). Weakly Supervised Multi-Task Representation Learning For Human Activity Analysis Using Wearables. *Proceedings Of The Acm On Interactive, Mobile, Wearable And Ubiquitous Technologies*, *4*(2), 1-18.

Eldele, E., Ragab, M., Chen, Z., Wu, M., Kwoh, C. K., & Li, X. (2024). Label-Efficient Time Series Representation Learning: A Review. *Ieee Transactions On Artificial Intelligence*.

Sheng, T., & Huber, M. (2019, October). Siamese Networks For Weakly Supervised Human Activity Recognition. In *2019 Ieee International Conference On Systems, Man And Cybernetics (Smc)* (Pp. 4069-4075). Ieee.

Eldele, E. A. I. A. (2023). Towards Robust And Label-Efficient Time Series Representation Learning.

Sheng, T., & Huber, M. (2020, May). Unsupervised Embedding Learning For Human Activity Recognition Using Wearable Sensor Data. In *Flairs* (Pp. 478-483).

Keating, L. (2024). Cross-Modal Weakly Supervised Learning For Multisensor Human Activity Recognition.

Rizk, H., & Elmogy, A. (2025). Self-Supervised Wifi-Based Identity Recognition In Multi-User Smart Environments. *Sensors*, *25*(10), 3108.

Sheng, T.; Huber, M. Consistency Based Weakly Self-Supervised Learning For Human Activity Recognition With Wearables. In Proceedings Of The Aaai-22 Workshop

On 4. Human-Centric Self-Supervised Learning, Virtual, 22 February–1 March 2022. Hc-Ssl'22.

Trirat, P., Shin, Y., Kang, J., Nam, Y., Na, J., Bae, M., ... & Lee, J. G. (2024). Universal Time-Series Representation Learning: A Survey. *Arxiv Preprint Arxiv:2401.03717*.

Sheng, T.; Huber, M. Reducing Label Dependency In Human Activity Recognition With Wearables: From Supervised Learning To Novel Weakly Self-Supervised Approaches. Sensors 2025, 25, 4032.

Surisetti, Vijaya. (2025). Ai-Driven Orchestration In Soa: Adaptive Workflows For Cloud- Based Enterprise Applications. International Journal Of Pharma Professional's Research (Ijppr). 6. 2868-2880.

Deldari, S. (2024). *Learning From Multimodal Time-Series Data With Minimal Supervision* (Doctoral Dissertation, Rmit University).

Liu, C., Gui, G., Wang, Y., Ohtsuki, T., Niyato, D., & Shen, X. S. (2025). A Comprehensive Survey On Self-Supervised Learning For Specific Emitter Identification. *Ieee Communications Surveys & Tutorials*.

Ding, C., & Wu, C. (2024). Self-Supervised Learning For Biomedical Signal Processing: A Systematic Review On Ecg And Ppg Signals. *Medrxiv*, 2024-09.

Azizi, S., Culp, L., Freyberg, J., Mustafa, B., Baur, S., Kornblith, S., ... & Natarajan, V. (2022). Robust And Efficient Medical Imaging With Self-Supervision. *Arxiv Preprint Arxiv:2205.09723*.

Xu, M. A., Narayanswamy, G., Ayush, K., Spathis, D., Liao, S., Tailor, S. A., ... & Mcduff, D. (2025). Lsm-2: Learning From Incomplete Wearable Sensor Data. *Arxiv Preprint Arxiv:2506.05321*.

Himeur, Y., Varlamis, I., Kheddar, H., Amira, A., Atalla, S., Singh, Y., ... & Mansoor, W. (2023). Federated Learning For Computer Vision. *Arxiv Preprint Arxiv:2308.13558*.

Wichrowski, F., Ostrowski, M., Bortyn, M., & Kaczmarek-Majer, K. (2025). Review Of Explainable Semi-Supervised Methods In Multivariate Time Series Analysis. *Authorea Preprints*.

Olivia, S. (2025). Weakly And Self-Supervised Learning Strategies For Human Activity Recognition Using Wearable Sensor Data.