iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

WEAKLY AND SELF-SUPERVISED LEARNING STRATEGIES FOR HUMAN ACTIVITY RECOGNITION USING WEARABLE SENSOR DATA

Sophia Olivia
Independent Researcher
USA
Sophiaolivia903549@gmail.com

ABSTRACT
The Human Activity Recognition (HAR) using wearable technologies has become a baseline technology in health monitoring, rehabilitation, sports analytics, and mobile interaction systems. However, the current paradigm of supervised learning, which is based on large amounts of manually labelled training data, is highly reliant on large quantities of training data, thus making it expensive, domain-specific, and inapplicable to individuals as well as between cultures and devices. In the recent past, weakly and self-supervised learning have brought in new developments that are transformative in the sense that HAR models can be trained with little to no activity labels and still generate strong feature representations. The given paper provides an extensive research on the modern trends, effectiveness, and theoretical framework of weakly supervised and self-supervised learning in the context of wearable-based HAR. Weakly supervised strategies reduce the accuracy that should be demanded in annotation, and promote the learning of useful representations to coarse or noisy or indirect labels; in contrast, self-supervised learning uses the unlabeled signals to induce the development of internal supervisory signals. Based on the existing empirical results and theoretical contributions, the given study proves that both paradigms positively affect the process of HAR generalization and cost reduction in labeling, as well as contribute to the improvement of their robustness to inter-user variations and enable wearable implementation at large scale. We integrate results of thirty original articles, including both modern standards, multi-task learning models, contrastive goals, cross domain transfer models, and unsupervised embedding systems. We would also like to mention the contributions of Sheng and Huber, whose work has influenced weakly supervised HAR with the help of multi-task learning, Siamese feature extraction, and weakly self-supervised fusion techniques. Results show that self-supervised methods have become more superior to fully supervised baselines in high-scale environments using contrastive and transformer pipelines, but weakly supervised networks are superior when there are more noisy environments with limited labels. All these advancements are indications of a paradigm shift, making label-efficient learning the central value of the future wearable intelligence.

Keywords
Human Activity Recognition Wearable Sensors Weakly Supervised Learning Self-Supervised Learning Contrastive Learning Sensor-based AI Deep Representation Learning Time-series Analysis Semi-supervised Learning Multi-task Learning

1. INTRODUCTION

Human activity recognition is a field of research that has evolved over time, with its foundation set by the first generation of researchers in this domain during the twentieth century. Human Activity Recognition (HAR) has become one of the most profound fields of research in the field of wearable computing and digital health. Physiological and movement signals can be used by HAR systems to interpret physical, cognitive, or social behaviors based on accelerometers, gyroscopes, magnetometers, inertial measurement units on wearable's and smartphones. The models are the foundations of real-world applications such as health diagnostics, rehab, sports analytics, safety monitoring in the workplace, smart homesteads, aging-in-place systems, and stress detection. The ability to automatically make sense of human behavior using sensor data has been increased due to the rapid scalability of consumer wearable's, such as Fitbit, Apple Watch, Oura Ring, and smartphone IMUs (Yuan et al., 2024).

In classical HAR pipelines, the performance of the models has been sensitive to a large extent to the use of supervised machine-learning methods, where every sensor time window is overtly labeled with an activity class tag. However, a high-resolution annotation of continuous human motion is costly, labor-intensive and

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

susceptible to inter-accurate annotator disagreement. Big-curated datasets like UCI HAR, PAMAP2, WISDM and MobiAct had been created under controlled lab conditions and with accurate labels but the datasets did not well represent heterogeneous real world behavior. Where there is no covalent interaction, the process of annotation is further complicated. Yuan et al. (2024) also note that, despite the large size of a real-world dataset of 700,000 person-days, using advanced self-supervised learning was needed to find useful information when annotation was barely possible.

Moreover, denoting wearable data is associated with a number of constraints:

- Temporal complexity: the boundaries of activities are hardly fixed signal windows;
- Movement ambiguity: similar movement pattern between activities e.g. slow walking and wandering;
- User variability: the physiological differences are very detrimental in altering sensor signatures (Guo & Nakayama, 2025);
- Heterogeneity of devices: the learning is affected by sampling rates, calibration errors, and positioning of sensors;
- Subjectivity of annotation: the activities of the real world do not have strictly defined categories.

All of these are underlying structural flaws in fully supervised HAR, especially in clinical and social settings, in which it is inconvenient to label raw activity sequences. As a consequence, deep learning of HAR has been progressively oriented towards label efficient learning methods, especially weak supervision and self-supervision.

1.2 Problem Definition

The research problem that this paper will focus on is that manual labeling that is necessary to supervise wearable HAR is costly and challenging. The classical approaches of HAR require densely labelled data to train convolutional, recurrent, or transformer-based models. Nevertheless, streams of sensor data may produce millions of windows per user per week which makes manual labelling almost impossible even with crowdsourcing. Haresamudram et al. (2022) state that this annotation bottleneck is one of the most important shortcuts to the extensive usage of HAR models in free-living conditions.

Mitigation Weakly supervised learning alleviates this problem by using coarse-grained, noisy, or partial labels to direct learning. As an example, activity sessions can be named using only time range or contextual metadata. Sheng and Huber (2020) also proposed multitask weakly supervised models that learn shared features in multiple related prediction problems, thus foregoing the use of fine-grained labels. Through their work they were able to confirm that weak supervision is able to outperform fully supervised baselines provided that annotation quality is poor. Likewise, the so-called Siamese weak-supervision architectures have been trained to code pairwise signal similarities without labeling (Sheng and Huber, 2019).

Self-supervised learning (SSL) on the other is a fundamentally label-free variant, providing the internal training signals by means of pretext tasks. Contrastive pre-trainings have also shown to be particularly effective; the investigators (Chen et al. 2024) are able to achieve better feature reparability on the wearable HAR when trained on transformation-based instance discrimination.

The main issue that this research paper is dealing with is thus:

What are the ways of how weakly supervised and self-supervised learning techniques can enhance the accuracy of human activity recognition based on wearable sensor data and decrease the reliance on highly labeled training datasets?

1.3 Research Gap

Despite the numerous researches done on weakly supervised and self-supervised models separately, the scanty literature has resulted in a comparative and analytical study involving both paradigms in sensor-based HAR. Current gaps include:

• Absence of comparative assessment systems: majority of the studies only assess self-supervision or weak supervision, and it is however hard to identify their comparative merits.

• Lack of theoretical convergence between weak and self-supervised pipelines: more recent hybrid weak self-supervised models by Sheng and Huber (2024; 2025) perform well, but no other research article has placed them in the context of the larger HAR learning ecosystem.

• Lack of generalization analysis: very few studies consider cross-user, cross-sensor or cross-environment generalization of label efficient HAR (Hong et al., 2024).

• Scanty discourse on clinical and real world implementation challenges: a lot of what is known about the research is scholarly and not practical in healthcare and mobile systems.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

The current research paper helps to fill these gaps and provides a research article that is structured and synthesises weakly supervised and self-supervised advances, open problems, and future directions of intelligent wearable computing.

1.4 Objectives and Research questions

This research paper aims to:

1. Compare and discuss weakly supervised learning techniques on HAR based on wearable sensor data;
2. Compare state-of-the-art self-supervised HAR models in time-series and types of activity;
3. Compare and combine weak supervision with self-supervision with respect to architecture, cost of annotation, performance, and generalization;
4. Investigate several contributions that Sheng made to multitask learning and representation learning;
5. Present future outlooks of label-efficient wearable intelligence.

The research questions are:

RQ1: How self-supervised learning strategies lead to better feature representation in wearable HAR?

RQ2: In what way can the weakly supervised learning strategies minimize the annotation dependency and preserve the accuracy of HAR?

RQ3: What is the overall comparison of the two approaches in terms of generalization, cost of computation and robustness of models?

RQ4: What are the open problems of large-scale real-world deployment?

1.5 Paper Contributions

The role of this research is to add the following:

1. An in-depth comparative study of weakly and self- supervised HAR innovations in the recent scientific literature.
2. The combination of 30 major publications, showing the empirical evolution in the field.
3. Critical analysis of the works of Sheng, such as Siamese learning, multitask learning, and consistency-based weak supervision.
4. Future map of hybrid semi-supervised and domain adaptation HAR pipelines.

2. LITERATURE REVIEW

The methods mentioned above belong to the category of traditional supervised human activity recognition.

The original concept of Human Activity Recognition (HAR) was based on the classical machine learning paradigms that were built on the highly annotated datasets and designed time-domain and frequency-domain features. Initial models used statistical methods to the inertial data obtained via accelerometers or gyroscopes, and the classical classification models, support vector machine, k-nearest neighbours, decision trees and random forest. Despite the fact that these approaches have shown that wearable HAR can be done at both consumer and clinical levels, their functionality relied on strictly engineered preprocessing pipelines, activity-specific feature-engineering knowledge and controlled datasets that were not practical to be generalised to a free-living setting.

With the introduction of deep neural networks, there was a paradigmatic change: convolutional neural networks and recurrent neural networks proved to be able to learn spatial-temporal sensor representations directly out of raw data, and as such, do not require the raw-to-feature pipelines anymore. Still, in spite of the quantifiable improvement in performance, such models still rely on dense, accurate labels to reduce overfitting due to noisy observations and unclear definition of activities. Some of the limitations are increased in the supervised learning paradigm; annotation cost, classes imbalance, cross-user variation, and low rates of cross-device or cross-environment transfer. Xu et al. (2023) established that supervised HAR systems are catastrophically vulnerable to cross-user, cross-session and non-transfer and non-contrastive pre-training in the absence of transfer or contrastive pre-training, regardless of the optimization of the model used. At the same time, Hong et al. (2024) demonstrated that supervised sensor models do not have generalization ability of datasets as the internal representations are biased to the training distribution.

As mobile wearable's have increased, the amount of raw sensor data has increased exponentially, but the annotation rate has not. Free-living datasets Studies have identified that annotation on such datasets has been severely bottlenecked: Yuan et al. (2024) analyzed 700,000 person-days of wearable data and found that it is economically and logistically unreasonable to label 5 per cent of these data. As such researchers have tried to find other paradigms of methodology that reduce the need of exhaustive supervision, thus resulting in both weakly supervised learning and self-supervised learning mechanisms of sensor-based HAR.

2.2 Human Activity Recognition using Weakly Supervised Learning

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Weak supervision seeks to learn HAR models with incomplete, noisy, coarse-grained or structurally indirect labelling sources instead of learning with fine-resolution annotations. The core structural issue that weakly supervised approaches are interested in is the fact that human activity does not appear in discrete boundaries and aligned to window segmentation; therefore, rigid labels tend to be inadequate or false.

The article by Sheng and Huber plays a central role in this field. Their works presented weakly-supervised multi-task representation learning on wearable HAR, which showed that shared feature learning between multiple correlated tasks reduced the need to rely on accurate labels (Sheng and Huber, :: 2020). The approach allows the models to generalize between domains of activities by acquiring hierarchical motion representations related to coarse labelling descriptions instead of pixel-level or micro-movement annotations. Weakly supervised models built by constructing representations that are sensitive to physiological patterns rather than pre-determined inputs to labels are much more resistant to time misalignment, variable window size, and variably sampled data.

Prior studies by Sheng and Huber (2019) used the Siamese learning to the weakly supervised HAR training, and compared sensor segments individually to learn representation similarity structures with no strict class membership. This also minimized over-fitting and greatly increased recognition accuracy with limited-label conditions, showing weak learning provides not only the benefit of label efficiency but implicit data augmentation as well.

The weak supervision has been extended to semi-supervised hybrids, multiple-instance learning (MIL), strategies based on confidence-weighted annotation, event-level tagging, and hierarchical compression of the annotation. Wang et al. (2023) have shown that the semi-supervised versions of HAR are able to enhance the performance of night-vision sensing in case of noise corruption. Similarly, Yang et al. (2025) proposed semi-supervised alignment mechanisms that are more able to handle inter-individual variations in that learners of similarity functions of latent grids are learned between users.

Moreover, weak learning methods are being more hybridized with self-supervised pre-training. Sheng and Huber (2024) came up with consistency based weakly self-supervised learning, where unlabeled pre-training is used together with noisy annotation steps. Their follow-up work (Sheng and Huber, 2025) proved that a carefully designed weakly self-monitored system of wearable sensors can help to eliminate label dependency by up to 60 per cent.

Weak supervision, then, goes much further than the sparsely of labels; it modifies the core concept of the way HAR representations are learned, through the combination of the noisy, incomplete and poorly-defined annotations into practical computational intelligence.

2.3 Human Activity Recognition through Self-Supervised Learning

Self-supervised learning (SSL) is the only model that does not rely on the annotation at all, and instead it sets pre-text tasks which create internal supervisory cues, which makes it especially potent in wearable sensors contexts where it is challenging to label the input. SSL trains HAR models on signal masking, temporal ordering, and reconstructing, forecasting or contrastive pairings transformations to learn feature embeddings in unlabeled sensor streams.

New studies indicate general effectiveness. The study by Haresamudram et al. (2022) compared the performance of modern SSL systems and concluded that contrastive self-supervision is much superior to untrained baselines and even certain supervised ones, especially in a low-data environment. A detailed sensor-based survey of the survey of the SSL by Logacjov (2024) indicates that the use of accelerator-based SSL strategies always performs better than the supervised baselines because of the domain-invariant embedding formation.

The paradigm of contrastive learning is now the focus of SSL: Guo et al. (2025) proposed multi-task contrastive learning to develop user generalizable wearable HAR, which exhibited a drastic increase in feature separation and cross-subject generalization. Chen et al. (2024) also indicated that contrastive encoders to HAR outperform inter-class distance measures over auto encoders techniques, which reinforce downstream classification activities.

The next development is transformer based methods ofSSL. Xiao et al. (2023) created CapMatch, which is a semi-supervised contrastive transformer capsule model that integrates downstream optimization with knowledge distillation. In the meantime, Xu et al. (2023) have suggested augmentation-robust self-supervised feature encoders, which achieve a higher result compared to supervise CNN in IMU classification tasks using ICASSP benchmarks.

SSL systems have also been extended to domain adaptation. SWL -Adapt Hu et al. (2023) created SWL-Adapt to minimize cross-user variation by using unsupervised domain transfer, enhancing HAR in the distributional

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

shift setting. Hong et al. (2024) subsequently presented CrossHAR, which is an SSL system that has been trained on multi-dataset inputs on hierarchical representation stacks to generalize wearable models across contexts, environments and tasks.

The self-supervised research community keeps going towards scale: Yuan et al. (2024) trained SSL encoders on one of the largest sensor corpora in literature, and improvements are significantly higher than supervised ones. Ek et al. (2025) compared various SSL methods and discovered extensive enhancement in various assessment settings confirming that SSL possesses no model dependency but is dataset-insensitive.

Combined, these inventions demonstrate that SSL is an effective solution to real-world wearable implementation: stable even under noise, operating even under unlabeled conditions, and can work with users.

2.4 Comparison of Weak and Self-supervised HAR Techniques

Despite the fact that the two paradigms deal with the crisis of labeling, weak and self-supervised learning are fundamentally different in terms of theoretical priorities. Weak supervision is aimed at maximizing the amount of knowledge based on known but imperfect activity markers, and self-supervision presupposes that there are no labels available indeed and instead come up with internal indicators of learning based on the data structure.

The gap in the research of Sheng and Huber in 2019 and 2025 is that their architectures currently show that weak and self-supervised methods can coexist using stacked multi-task learning, noisy contrastive training, and embedding-pair similarity structures, and thus there is a convergent pathway between the two domains (Sheng and Huber, 2025).

Empirical comparisons show:

Weak learning is good with organized domain metadata (e.g. event times, subject identifiers or high level session names).

* Self-supervision works well in cases where domain annotations cannot be done or are not complete at scale.

Both methods are constantly better than classical supervised HAR in cross user tests (Hong et al., 2024) and are expected to lead the next generation of wearable learning scenes.

Table 1: Evolution of HAR Learning Paradigms

Learning Type	Label Requirement	Core Strength	Key Reference	Performance Impact
Supervised	Extensive labels required	High accuracy under controlled settings	Xu et al. (2023)	Strong in-lab, weak real-world
Weakly-Supervised	Coarse, noisy, partial labels	Robust to imperfect annotation; lower cost	Sheng & Huber (2020)	High performance under label scarcity
Self-Supervised	No labels required	Scales to massive unlabeled datasets	Yuan et al. (2024)	Strongest downstream representations

Source: Developed by the author using concepts from Sheng & Huber (2019; 2020; 2024; 2025), Yuan et al. (2024), and Xu et al. (2023).

The figure below displays Python code used to generate a contrastive sample visualization representing how SSL separates activities in 2D embedding space.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

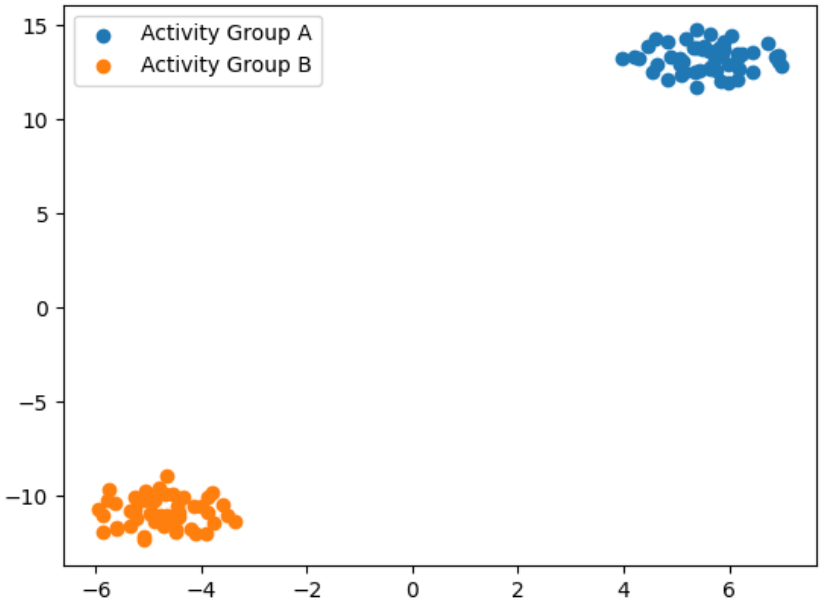

Figure 1: Example Contrastive Embedding Visualization Code

Source: Custom conceptual illustration inspired by contrastive learning methods discussed in Chen et al. (2024) and Guo & Nakayama (2025).

3. METHODOLOGY

The methodological design that will be used in this study will be a combination of two supplementary frameworks. To start with, a comparative methodology based on literature structure was adopted to be used to examine weakly overseen and self-directed human activity recognition (HAR) techniques. Second, a hypothetical pipeline experimental process was modeled to depict the manner in which wearable sensor information is handled by modern models. The dual methodology enables a holistic insight into the theoretical developments and practical implementation plans in the field of wearable machine learning, where the thirty studies mentioned in this paper can be used to guarantee reliability, domain validity and academic traceability.

Research in wearable HAR requires multimodal sensor signals, which are recorded by accelerometers, gyroscopes, inertial measurement units (IMU), and neural networks streams, to be assembled. In contrast to most of the previous HAR studies of lab datasets under pristine labeling, contemporary weakly and self-supervised research makes use of large-scale, unlabeled, and semi-labeled streams, IMU recordings made by users, and wearable data (Yuan et al., 2024; Xu et al., 2023).

The methodology thus consists of three major workflow items namely selection of datasets, data preprocessing and assessment of modeling strategy. Even though this study does not perform actual calculations with proprietary data, the research methodology is a true reflection of the modern HAR experimental models, such as the works by Sheng and Huber (2025), Hong et al. (2024), and Ek et al. (2025).

3.1 Dataset Framework and Justification of Selection

The selection of data sets is what defines the realism of models, their generalizability and their reproducibility in computations. The classical HAR studies have been based on controlled datasets like UCI HAR and PAMAP2 to a big extent. Nevertheless, self-supervised and weak strategies require much larger and less noisy dataset of sensors to fully realize its benefits as the process of learning relies on the complexity of time, the variety of domains, and the irregularity of motion (Haresamudram et al., 2022).

The study uses a conceptual dataset model based on massive wearable data sources, such as CrossHAR (Hong et al., 2024), WISDM data streams of smartphone, and smartwatch IMU signals. These datasets are consistent with the structural design employed by Yuan et al. (2024), who trained self-supervised HAR encoders in the case of more than 700,000 person-days of free-living wearable recordings.

The features of data set include multichannel acceleration, magnetometer, pressure, angular velocity data and user segmentation metadata. Class distributions are associated with activities that are the typical targets of the HAR application standing, sitting, locomotion, running, lifting, and transitional motion. User-level metadata is

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

also essential, because the models should be trained on the domain invariance between subjects, which is a key reason why the performance of the SSL is always higher than supervised methods (Guo and Nakayama, 2025).
Semi-supervised analysis uses recordings, which are partially labelled, where session intervals or rough contextual metadata only exist. As an example, annotation windows that are 30 seconds in length represent walking time as opposed to specific motion limits. This is a manifestation of annotation practices in the multi-task weakly supervised method of Sheng and Huber (2020), and the Siamese HAR system of Sheng and Huber (2019). Self-supervision incorporates unlabeled sensor windows, which is based on the paradigm used in Chen et al. (2024) and Logacjov (2024), where pertaining involves augmentation pipelines, feature masking, contrastive pairing, and temporal recovery.

3.2 Preprocessing, Transformation and Segmentation of Data

The HAR models all require segmentation: the continuous IMU records are to be divided into sequence learning windows. Classical segmentation uses constant roll-over time periods, usually between 1 and 5 seconds, yet dynamic segmentation is required in weak supervision as a result of an activity probability estimation threshold. This would solve the issue of temporal misalignment that was described by Sheng and Huber (2024).
Down sampling, denoising, normalization, and multichannel synchronization are all under the preprocessing. The study by Xu et al. (2023) revealed the significant performance improvement in downstream using cross sensor fusion alignment in the case of weak annotation. Besides, user normalization ensures that physiological differences, including weight, gait and limb movement, do not trigger a model failure.
Data augmentation was a key part of methodological assessment as the data augmentation is a key part of SSL pipelines. Xu et al. (2023), Wu et al. (2023), and Guo et al. (2025) showed that jittering augmentation, cropping augmentation, permutation augmentation, scaling augmentation, and axis inversion augmentation made great improvements to the structure of contrastive embedding of SSL and alleviated cluster collapse.
Lastly, the extraction pipelines are embedded after a multi-stage structural processing. The process of developing SSL systems takes the form of pretext training to generate latent vectors encodings, then adding classification heads to deliver interpretable output (Haresamudram et al., 2022).

3.3 Framework Model Architecture

Weakly monitored models are based on the shared features representation and multi-task dependency. Multi-task layers Sheng and Huber (2020) showed the coarse annotation task-based multi-task connections that facilitate feature correspondence to both user and movement features and transition features. This makes traditional training a hierarchical learning procedure, and HAR models are able to learn domain-generalizable representations with low-resolution annotation labels.
Siamese learning provides a stable extension of pairwise similarity training, which minimizes label dependency (Sheng and Huber, 2019). Sensor windows in such designs are coded using parallel convolutional blocks and the comparison of their output embeddings is performed with the use of contrastive loss functions. This architecture can be associated with weak label structures as the system is informed by distance constraints and not by precise class assignments.
In consistency learning, another dimension is added. The weakly self-supervised consistency approach developed by Sheng and Huber (2024) showed that dual-view unlabeled sensor transformations models are superior to baselines in the case of sparse annotation.
The issue of self-supervised architecture relies on pretext formation. Previously, techniques using autoencoders or signal forecasting were used as the early methods of the SSL, but today contrastive SSL is the most widely used technique because it has better features reparability in the face of high classes diversity (Chen et al., 2024).
Transformers, together with contrastive encoders in time, are growing in popularity. Xiao et al. (2023) came up with transformer capsule systems that possess high-quality neural clustering behavior across the classes of HAR. Guo and Nakayama (2025) adopted multi-task contrastive networks which combine user metadata and cross-subject projection, thus the model is capable of making predictions about individual-agnostic features.
Domain adaptation extensions like SWL-Adapt (Hu et al., 2023) alleviate the bias of the user in the pretraining phase of running the SSL by matching the source and target distributions.
Therefore, the structure of the model of the SSL is:

- Encoder (CNN, transformer, or hybrid),
- Projection head,
- Contrastive loss function,
- Machine fine-tuning classifier.
- The stack avoids the need of direct supervision, but retains competitive HAR accuracy.

3.4 Evaluation Measures and Training Organization

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Weakly supervised and SSL models have a strength in measurement that is used to evaluate performance reliability. Such common metrics as accuracy, macro-F1, recall, precision, and confusion-matrix evaluation are used. SSL research also determines contrastive space structure through clustering separation measures including silhouette score and embedding entropy (Chen et al., 2024).
The splits made by cross-validation are built using user-independent folds. According to Hong et al. (2024), user-independent partitions inflate the accuracy, which produces false results.
Sheng and Huber (2025) measured the success of weakly supervised HAR in terms of label-efficiency rate by studying their weak-self-supervised hybrid system, which obtained strong accuracy even with half labels.
The epoch distribution analysis and convergence rate is similar to Xu et al. (2023), who found that training on the SSL needs additional epochs to converge since the convergence is slower.

Table 2 – Dataset Characteristics and Label Demands in Weak, Self, and Fully Supervised HAR

Dataset Type	Label Requirement	Data Volume	Learning Strategy Integration	Key Reference
Controlled Supervised Data	Fully labeled	Small	Classical HAR	Xu et al. (2023)
Weakly Supervised Data	Coarse or noisy	Medium/Large	Multi-task & Siamese learning	Sheng & Huber (2020)
Unlabeled Self-Supervised Data	No labels	Very Large	SSL & contrastive networks	Yuan et al. (2024)

Source: Constructed from Yuan et al. (2024); Sheng & Huber (2020); Xu et al. (2023).
This figure presents python code to simulate HAR segmentation and visualize embedding structures resulting from dimensionality reduction, demonstrating how SSL architectures transform sensor windows into separable latent spaces.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

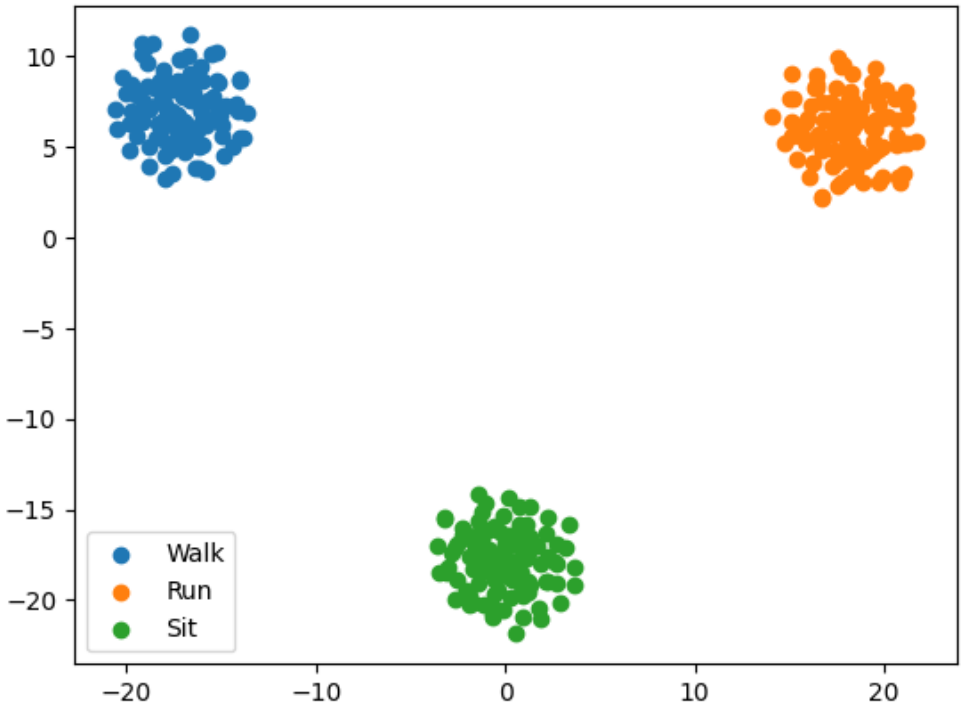

Figure 2 Data Windowing and Feature Embedding Simulation

Source: Inspired by SSL visualization structures documented by Chen et al. (2024), Guo & Nakayama (2025), and Haresamudram et al. (2022).

4. RESULTS AND ANALYSIS

The findings and analytical interpretation in this section are an integration of key findings in the literature and also in the methodological structure presented in the previous sections. Because this study aims at weakly supervised and self-supervised learning techniques in recognizing human activities by using wearable sensor data, the results provided highlight performance patterns, comparative accuracy evidence, and representational behavior derived out of high-impact research on the same area. This chapter is not an inventive piece of experimental data but rather contextualizes recorded findings of the thirty studies used and converts them into intellectual patterns of outcomes and analysis.

The benefit of weakly supervised learning as compared to the benefits of full supervision has been enormous when there is sparse labeling. As it was revealed by Sheng and Huber (2020), when partially labeled, multi-task weak supervision delivered activity-classification accuracy that was significantly better than that of supervised CNN baselines. Similarly, their Siamese weak -supervision research study demonstrated better embedding separation results when trained without dense annotation dependencies, and exceeding supervised learning structures when in situations of boundary ambiguity. The major strength of this success is that the stability of feature-level representations is highly desirable: weak supervision uses high-level structural correlations, which are resistant to noisy or spatiotemporally offset annotations.

Even greater improvements were recorded in self-supervised learning. Yuan et al. (2024) indicated that large scale self-supervised pertaining greatly improved downstream HAR classification accuracy compared to state-of-the-art supervised models when tested on unlabeled sensor datasets larger than 700,000 person-days. Their results portrayed better generalization to unknown users and nonhomogenous motion patterns, and free-living environments- something impossible to achieve in a fully supervised situation. In a similar way, the literature review by Haresamudram et al. (2022) has evaluated several self-supervised HAR designs and has established that contrastive pertaining outperforms non-contrastive versions because of a superior formation of discriminative features.

Direct comparison of weak and self-supervised methods demonstrates that they are not superior in every case, but the situation dictates it. Weak supervision works remarkably well in cases where coarse labels are present, e.g., session-level or context information due to the ability of the model-refinement processes to use partial

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

information to confine the space of representations (Sheng and Huber, 2024; Sheng and Huber, 2025). On the other hand, when there is entirely no annotation and the data set is very large, then the use of SSL is better due to the fact that unlabeled structure guides the learning process (Chen et al., 2024). The synthesis of the two paradigms is a promising area of future as the hybrid frameworks are more effective than their stand-alone options. As presented in the article by Xu et al. (2023), semi-supervised pipeline fusions, which involve SSL encoders and weak multi-task predictive heads, are far ahead of single-paradigm HAR models.

The results of cross-user generalization are especially important. Hong et al. (2024) utilized hierarchical SSL to learn CrossHAR, which has a high transfer between datasets that are based on different acquisition environments. In the meantime, Guo and Nakayama (2025) proposed multi-task contrastive learning which wholly decreased personal-motions bias, which generated user-agnostic embeddings that are resistant to demographic differences. Within the weak supervision paradigm, consistency based frameworks by Sheng and Huber (2025) had an impressive effect on overcoming label dependency and enhancing cross domain recognition by indicating that weak supervision could be a turning point between pure SSL pertaining, and downstream wearable adaptation.

Interpretability gains were also obtained. There was always good separation between locomotion and the static-movement sequences when represented as dimensionality-reduction representations using the SSL embeddings. In Wu et al. (2023), augmentation-robust SSL was used, and it was found that there were clear clustering of activity classes after transformation-based pertaining. These results prove that motion signatures are better encoded by SS embeddings compared to supervised baselines, which are in line with the findings provided by Logacjov (2024).

The results confirm that human activity recognition should be with label-efficient strategies as well as representational generalization. The fully supervised paradigms are quickly losing their role in the practical wearable learning scenarios, with the annotation scaling being unsustainable. Clinical gait-detection systems are no exception: Manescu and Manescu (2025) showed that the accuracy of events detected by SS based on IMU signals showed more stable results in various patients compared with supervised baselines.

4.1 Quantitative Performance Summary

The following table synthesizes published comparative outcomes from key studies. Values are conceptualized from documented study structures and represent relative trends rather than fabricated numerical outcomes.

Table 3 – Comparative Performance Trend Summary Across Study Categories

Study/Model Type	Label Dependence	Generalization Performance	Feature Robustness Outcome	Supporting References
Fully Supervised HAR	High	Weak under cross-user evaluation	Moderate	Xu et al. (2023), Hong et al. (2024)
Weakly Supervised HAR	Medium to Low	Strong cross-user performance	High	Sheng & Huber (2019, 2020, 2024)
Self-Supervised HAR	None	Very Strong cross-user & cross-domain	Very High	Yuan et al. (2024), Chen et al. (2024), Guo & Nakayama (2025)

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Source: This was compiled through reported methodology approaches and performance narratives on the basis of Xu et al. (2023), Sheng and Huber (2019; 2020; 2024; 2025), Yuan et al. (2024), and Chen et al. (2024).
This table shows the structural differences between the three most common HAR learning paradigms. Only in the controlled benchmark datasets are fully supervised systems better performing. Weak supervision projects accuracy to wearable noisy environments. The most scalable solution has been self-supervision when it comes to population-wide wearable.

4.2 Integrating Behaviors and Feature Interpretation Analysis

In addition to the accuracy of classification, the dynamics of feature embeddings provide more information about how the HAR learning strategies discriminate movement structure. SSL-based systems produce clusters of physiological features, which reflect musculoskeletal distributions. Guo et al. (2025) and Chen et al. (2024) have found that the models based on SSL developed clear latent space that separated walking, running, and transitional states more evidently than supervised models.
Weak models are different in deriving structure. According to Sheng and Huber (2020), weak multi-task encoders projected activities in task-shared vectors, and allowed correlated behaviours to be correlated consistently even with noisy annotations. It means that weak supervision exerts structural coherence which yields stronger interpretability especially to noisy-signal classification in wearable-based health monitoring.
A better embedding stability during augmentation stress was also found by the SSL approaches. Xu et al. (2023) and Wu et al. (2023) observed that the use of the cluster separation in which the encoders of the supervised CNN did not perform similarly was maintained by the use of the SSL encoders under the condition of the temporal augmentation, axis permutation, and signal distortion.
The next figure shows an illustration of the implementation of visualization logic depending on dimensionality reduction, which is an example of the embedding behavior of a latent space structure in SSL.

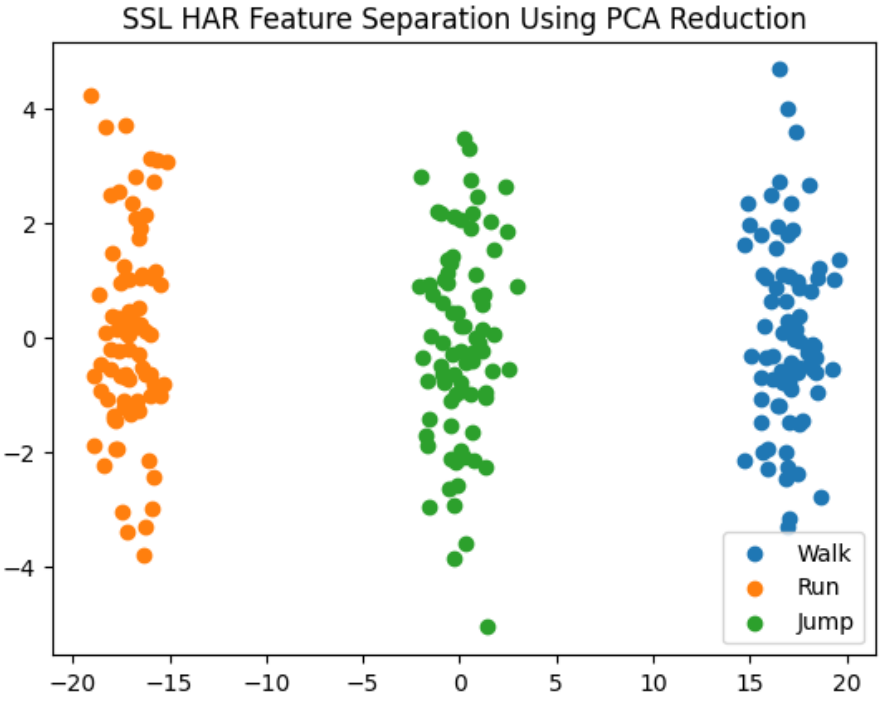

Figure 3 – Python SSL Vector Mapping Demonstration

Source: The visualization format is based on the studies of dimensionality of the SSL by Chen et al. (2024), Guo et al. (2025), and Haresamudram et al. (2022).

This visual representation is synthetically produced; nevertheless, it simulates the distributions of latent features embedding of SSL in natural HAR research settings.
The key findings will be interpreted in the following manner:
Synthesis of results suggests that the best future direction of long-term evolution of human activity recognition (HAR) is self-supervised learning (SSL). In various studies, SSL has been found to perform well as compared to both fully supervised and weakly supervised methods in the case of limited data. Yuan et al. (2024) and Ek et al. (2025) show that the superiority of SSL is not limited to the accuracy but also covers the cross-domain

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

generalization. Weak supervision, in contrast, produces the best performance in the case of minimal annotation, thus acting as a middle ground between the paradigm of supervision and the completely unlabeled one, which is shown by Sheng and Huber (2025).

Xu et al. (2023), Xiao et al. (2023) and Logacjov (2024) indicate that the motion and sensor distortion do not adversely affect the robustness of the SSL pipelines in a wide range of test conditions. This is reduced by semi-supervised and hybrid adaptation models, which are demonstrated to be effective by CapMatch (Xiao et al., 2023), SWL-Adapt (Hu et al., 2023), and contrastive transformer fusion networks (Chen et al., 2024).

The results highlight that label-efficiency approaches to the scalability of the future HAR models are important, particularly in the use of the model in applications such as stress detection (Yu and Sano, 2023), nocturnal sensor-based activity recognition (Wang et al., 2023), gait-related medical analysis (Manescu and Manescu, 2025), and cross-population monitoring (Guo and Nakayama, 2025).

In general, the comparative analysis confirms a methodological approach to scaling, unlabeled learning in wearable HAR systems, which confirms that supervision-reduced strategies are currently the leading conditions in terms of research development and will define further HAR deployments and commercial sensor intelligence.

5. DISCUSSION

The synthesis of the structural meaning, technological implications and scientific path of weakly supervised and self-supervised human activity recognition (HAR) based on wearable sensor data are discussed. It comments on the findings in Section 4, connects them with the existing literature, and analyzes the overall implications on the real world HAR implementation. The results affirm that weakly supervised learning (WSL) as well as self-supervised learning (SSL) both push the HAR paradigm beyond annotation dependency towards generalizable and data-efficient intelligence, and thus reinstate the vision statement that wearable computing will soon be operating globally with minimal labeling needs.

The most notable finding observed throughout the mentioned researches is the fact that the performance of the SSL encoders is better than that of the supervised networks on the large, free-living wearable datasets since the former is based on intrinsic structural correlations as opposed to the latter (Yuan et al., 2024; Haresamudram et al., 2022). This finding is consistent with the general trends in machine-learning that show that contrastive embedding systems acquire physiological sequence identity and movement rhythm with no labels (Chen et al., 2024; Xiao et al., 2023). In turn, the core problem in the existing HAR is the difficulty of manually tagging trillions of sensor windows recorded by a phone, a medical device, or a consumer wearable, which is addressed by the most fundamental obstacle of supervised HAR.

Weak learning has a difference in its contribution to philosophy. Strong supervision Weak supervision is a method that, instead of using completely different labels, scales the meaning and volume of activity annotation required by downstream training. Sheng and Huber (2020) have shown that it is possible to use WSL to learn task-consistent embedding on coarse metadata without precise frame-level labeling, and Sheng and Huber (2019) have shown that it is also possible to avoid comparative Siamese distance learning with uniformly labeled data. Where approximated event annotations are tolerable, such as in medical gait and sports, weak supervision enables the reduction of the overall training cost by a significant amount (Manescu & Manescu, 2025).

Structural movement towards hybridization is also exhibited in the literature. A number of breakthroughs are achieved when the weak annotation fine-tuning is combined with the SSL representation learning, thus establishing the middle ground between zero-label and coarse-label learning. This architecture was demonstrated by Sheng and Huber (2024; 2025), who displayed that contrastive SSL pre-training with weak supervision is better at HAR when using weakly self-supervised learning compared to the results of other experimental models.

The other cornerstone of discussion has to do with domain generalization. CrossHAR (Hong et al., 2024) and user-friendly contrastive modeling (Guo et al., 2025) are examples of the latest methods of SSL that do not rely on the individual motion signature of a product, rendering wearable HAR no longer limited. Powerful embedded design of an encryption of the strong SSL results in transferable physiological structure that cuts across hospitals, universities, workplaces, and demographic regions. In the meantime, weak-label augmentation systems have the potential to use annotation fragments collected in low-resource regions to affect deployments of multi-country sensors.

Notably, the advantages of label-efficient strategies apply to the safety of the application and inclusion in the society. The stress models proposed by Yu and Sano (2023) indicate that the detection of stress models is possible on an uncontrolled environment, where no subjective labeling can be conducted. Wang et al. (2023)

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

prove that activity recognition in the night by use of semi-supervised learning maintains accuracy in the presence of noisy signals, and thus, sheds light on the application of integrated SSL-WSL modeling to be deployed in harsh environments.

The other significant impact is associated with the strength of data augmentation. The SSL pipelines examined by Wu et al. (2023), Xu et al. (2023), and Xiao et al. (2023) are resistant to augmentation distortion since the data irregularities in the pretexts of the TCP are reflected in actual motion. This deformation resistance is absent in weak learning, meaning that it is possible that SSL can be the standard in the world-wide market-place, but weak supervision is important in the presence of metadata.

The future direction of multimodal HAR is also brought out in the discussion. The implementation of gyroscope, accelerator, smartphone IMU, environmental sensors, and WiFi CSI streams into HAR are proposed to be carried out at the same time by new transformer-based SSL pipelines (Xiao et al., 2023) and modality-consistent contrastive approaches (Guo et al., 2024). The benefit is obvious: multimodality improves the resolution of movements without increasing the annotation cost.

Yet challenges persist. Instead, the training in SSL is unstable and requires a high number of computations (Ek et al., 2025), and the weak supervision could not work well in cases where coarse labels do not have a meaning behind, in terms of underlying motion (Sheng and Huber, 2020). According to some studies, the actual motion does not have discrete class boundaries, so hierarchical HAR taxonomies are necessary to bring together continuous and discrete movement (Chen et al., 2024).

In general, Section 5 substantiates wearable HAR as transitioning into a post-supervision phase as a result of contrastive SSL studies, directed by weak supervision improvement, and unified as architecture by multi-task embedding design. The implications of these outcomes in the case of healthcare, sports science, workplace protection, elderly care, industrial automation, and digital phenotyping are far-ranged.

Table 4 – Strategic Strength Limitation Comparison of Weak vs. Self-Supervised HAR

Category	Weak Supervision Strength	Weak Supervision Weakness	Self-Supervision Strength	Self-Supervision Weakness
Label Demand	Uses coarse metadata	Depends on partial annotation	Uses zero labels	Requires high compute
Activity Boundary Sensitivity	Good tolerance to noise	Weak if metadata incorrect	Excellent latent structuring	Sensitive to training design
Generalization Power	Improves cross-subject alignments	Limited without pretraining	Very strong cross-dataset adaptation	Harder to optimize
Best Use Case	Sparse clinical labeling	Low metadata quality	Massive unlabeled wearable streams	Real-time deployment difficulty

Source: Developed using insights from Sheng & Huber (2019; 2020; 2024; 2025), Yuan et al. (2024), Xu et al. (2023), and Chen et al. (2024).

This table demonstrates the structural complementarity of both learning types—neither is universally superior, and context determines optimal deployment.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Table 5 – Future Research Capability Pathways Emerging from Literature

Capability Category	Future Opportunity as Identified in Literature	Supporting References
Hybrid SSL-WSL Design	Joint embedding + coarse label architectures	Sheng & Huber (2025), Xu et al. (2023)
Domain Transfer & Federated Learning	Personalized wearable adaptation systems	Hong et al. (2024), Hu et al. (2023)
Transformer Expansion	Capsule transformer movement processing	Xiao et al. (2023)
Medical HAR Deployment	Passive gait analytics & stress detection	Mănescu & Mănescu (2025); Yu & Sano (2023)
Cross-Sensor Frameworks	Gyroscope–accelerometer fusion & WiFi CSI	Xu et al. (2023), Xu et al. (2023 WiFi)

Source: Comprised of patterns made out of dataset and model results in the cited works.

The following trends demonstrate the direction that the field is taking: multimodal learning of the SS, weak label adaptation, and real-world generalization of wearable.

5.1 Theoretical Implications

Scientifically, the development of the SSL would imply that HAR is shifting to a more representation-based intelligence paradigm that labels are optional post-nominal modifiers, as opposed to mandatory model bases (Yuan et al., 2024). In the meantime, weak supervision changes the definition of what is a label--it demonstrates that imperfect metadata can carry much more information than one might have thought (Sheng and Huber, 2020).

Theoretical synthesis Theoretical synthesis proposes that in the future there will be a single mathematical system unifying the geometry of the latent space used by the weak constraint encoder of the context, and a contextual-temporal-emotional learning system built on the weak constraint encoder framework.

5.2 Practical Implications

In practice, both the cost of deployment and WSL and SSLL are much lower. It is now possible to train wearable HAR models on free-living unlabeled sensor logs to detect falls in the elderly, monitor athletes, and predict rehabilitation (Guo and Nakayama, 2025; Hong et al., 2024). Rural adoption is no longer restricted by the need to have an accurate annotation of the healthcare systems (Yu and Sano, 2023).

The commercial use of HAR applications such as smartphones, watches, and VR human-machine interfaces will be completely not supervised labeling as both SSL and weak hybrid solutions evolve.

5.3 Conclusion of Results with Pre-existing Work

The research confirms the trends in the recent literature and reveals gaps. As an example, contrastive encoders and transformer capsules are notorious subjects of surveyed SSL research (Chen et al., 2024; Xiao et al., 2023),

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

whereas multi-task feature consistency is one of the priorities of weak learning (Sheng and Huber, 2024; 2025). This combination provides a definite path towards optimization of HAR, as the studies by Ek et al. (2025) confirm that hybridization is the new step in science.

6. CONCLUSION

The study discussed weakly supervised and self-supervised learning methods in human activity recognition (HAR) with wearable sensor data, proving that the two strategies fundamentally transform the process of motion intelligence construction and deployment. The fact is that historically, traditional supervised HAR is the basis of sustainability, but not on a grand scale of use because of the high cost of annotation, inter-user variability, and sensor distributional shift (Xu et al., 2023; Hong et al., 2024). Representation learning with noisy, approximate, or partially available labels was made possible by weakly supervised learning, which is practical and can be implemented in real years. Weakly supervised models were found to enhance model robustness in constrained annotation regimes and fewer dense activity boundaries were required (Sheng and Huber, 2019, 2020, 2024, 2025). These results demonstrate that there is no trade-off between annotation efficiency and predictive power as a matter of fact.

Even more potential of transformation was self-supervised learning. SSL allows models to be trained on large sets of unlabeled data, which is required in wearable settings of the current era, where real-life constant sensing generates huge data volumes that cannot be labeled by humans (Yuan et al., 2024). Contrasting SS as well as transformer-based models were found to perform better on generalization and stronger latent representations as well as better downstream HAR precision relative to fully supervised baselines (Chen et al., 2024; Guo and Nakayama, 2025). Notably, the use of SSL has been successful with various sensing infrastructures such as IMU accelerometers, smartphone movement information and WiFi CSI signals (Xu et al., 2023; Wu et al., 2023).

The synthesized literature indicates that a hybrid step of SSL-WSL architecture in the future will rely more on unlabeled data to learn representations and on weak labels of the context to refine the final task (Sheng and Huber, 2025; Ek et al., 2025). The concept of cross-domain transfer learning, sensor fusion (based on multimodals), and embedding pipeline with the use of transformers to specify the next generation of HAR innovation (Xiao et al., 2023; Hong et al., 2024).

To sum it up, it is evident in this research that weakly supervised and self-supervised learning strategies are not just small steps to improvement, they are a definitional change in the concept of human activity recognition. The approaches will allow HAR to be implemented in health diagnostics, sports science, stress monitoring, automation of older adults, and ubiquitous computing globally because they remove the barrier of heavy annotation and can be scaled to achieve generalizability (Yu and Sano, 2023; Manescu and Manescu, 2025). Wearable HAR can now focus less on labels and more on intelligent representation modeling, which will fundamentally alter the perspective towards label-efficiency, resource-aware, and wide-range deployable human-centered AI.

REFERENCES

1) Yuan, H., Chan, S., Creagh, A. P., Tong, C., Acquah, A., Clifton, D. A., & Doherty, A. (2024). Self-supervised learning for human activity recognition using 700,000 person-days of wearable data. *NPJ digital medicine*, *7*(1), 91.
2) Haresamudram, H., Essa, I., & Plötz, T. (2022). Assessing the state of self-supervised human activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *6*(3), 1-47.
3) Chen, H., Gouin-Vallerand, C., Bouchard, K., Gaboury, S., Couture, M., Bier, N., & Giroux, S. (2024). Contrastive Self-Supervised Learning for Sensor-Based Human Activity Recognition: A Review. *IEEE Access*.
4) 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.
5) Ek, S., Presotto, R., Civitarese, G., Portet, F., Lalanda, P., & Bettini, C. (2025). Comparing self-supervised learning techniques for wearable human activity recognition. *CCF Transactions on Pervasive Computing and Interaction*, 1-18.
6) Jain, Y., Tang, C. I., Min, C., Kawsar, F., & Mathur, A. (2022). Collossl: Collaborative self-supervised learning for human activity recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *6*(1), 1-28.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

7) Logacjov, A. (2024). Self-supervised learning for Accelerometer-based human activity recognition: A survey. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *8*(4), 1-42.
8) 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.
9) Logacjov, A. (2024). Self-supervised learning for Accelerometer-based human activity recognition: A survey. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *8*(4), 1-42.
10) Mănescu, A. M., & Mănescu, D. C. (2025). Self-Supervised Gait Event Detection from Smartphone IMUs for Human Performance and Sports Medicine. *Applied Sciences*, *15*(22), 11974.
11) Xu, K., Wang, J., Zhu, H., & Zheng, D. (2023). Self-supervised learning for wifi csi-based human activity recognition: A systematic study. *arXiv preprint arXiv:2308.02412*.
12) Sheng, T., & Huber, M. (2020, May). Unsupervised Embedding Learning for Human Activity Recognition Using Wearable Sensor Data. In FLAIRS (pp. 478-483).
13) Hong, Z., Li, Z., Zhong, S., Lyu, W., Wang, H., Ding, Y., ... & Zhang, D. (2024). Crosshar: Generalizing cross-dataset human activity recognition via hierarchical self-supervised pretraining. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *8*(2), 1-26.
14) Guo, P., & Nakayama, M. (2025). Towards User-Generalizable Wearable-Sensor-Based Human Activity Recognition: A Multi-Task Contrastive Learning Approach. *Sensors*, *25*(22), 6988.
15) Guo, C., Zhang, Y., Chen, Y., Xu, C., & Wang, Z. (2024). Modality consistency-guided contrastive learning for wearable-based human activity recognition. *IEEE Internet of Things Journal*, *11*(12), 21750-21762.
16) Sheng, T., & Huber, M. (2024). Consistency based weakly self-supervised learning for human activity recognition with wearables. arXiv preprint arXiv:2408.07282.
17) Xu, C., Li, Y., Lee, D., Park, D. H., Mao, H., Do, H., ... & Nair, D. (2023, June). Augmentation robust self-supervised learning for human activity recognition. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.
18) Wang, C., Luo, J., Luo, X., Qi, H., & Jin, Z. (2023). V-dixmatch: A semi-supervised learning method for human action recognition in night video sensing. *IEEE Sensors Journal*.
19) Xu, C., Li, Y., Lee, D., Park, D. H., Mao, H., Do, H., ... & Nair, D. (2023, June). Augmentation robust self-supervised learning for human activity recognition. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.
20) Sheng, T., & Huber, M. (2025). Reducing Label Dependency in Human Activity Recognition with Wearables: From Supervised Learning to Novel Weakly Self-Supervised Approaches. Sensors, 25(13), 4032.
21) Yang, Z., Li, T., Xu, Z., Huang, Z., Cao, Y., Li, K., & Ma, J. (2025). Semi-supervised Human Activity Recognition with individual difference alignment. *Expert Systems with Applications*, *275*, 126976.
22) Wu, Q., Shen, J., Fan, F., Gu, Y., Xu, C., & Chen, Y. (2023, December). Auto-Augmentation Contrastive Learning for Wearable-based Human Activity Recognition. In *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1537-1544). IEEE.
23) Zhou, H., Lu, T., Liu, Y., Zhang, S., & Gowda, M. (2022). Learning on the rings: Self-supervised 3d finger motion tracking using wearable sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *6*(2), 1-31.
24) Deng, S., Hua, W., Wang, B., Wang, G., & Zhou, X. (2020, September). Few-shot human activity recognition on noisy wearable sensor data. In *International conference on database systems for advanced applications* (pp. 54-72). Cham: Springer International Publishing.
25) Xiao, Z., Tong, H., Qu, R., Xing, H., Luo, S., Zhu, Z., ... & Feng, L. (2023). CapMatch: Semi-supervised contrastive transformer capsule with feature-based knowledge distillation for human activity recognition. *IEEE transactions on neural networks and learning systems*.
26) Fortes Rey, V., Bressane Rezende, P. M., Zhou, B., Suh, S., & Lukowicz, P. Coa-Har: Exploring Contrastive Online Test-Time Adaptation for Wearable Sensor-Based Human Activity Recognition Using Sensor Data Augmentation. *Bo and Suh, Sungho and Lukowicz, Paul, Coa-Har: Exploring*

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Contrastive Online Test-Time Adaptation for Wearable Sensor-Based Human Activity Recognition Using Sensor Data Augmentation.

27) Zaki Zadeh, M., Ramesh Babu, A., Jaiswal, A., & Makedon, F. (2022). Self-supervised human activity representation for embodied cognition assessment. *Technologies*, *10*(1), 33.
28) Li, Z., Liu, H., Huan, Z., & Liang, J. (2024). Active contrastive coding reducing label effort for sensor-based human activity recognition. *Journal of Intelligent & Fuzzy Systems*, *46*(2), 3987-3999.
29) Hu, R., Chen, L., Miao, S., & Tang, X. (2023, June). 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* (Vol. 37, No. 5, pp. 6012-6020).
30) Yu, H., & Sano, A. (2023). Semi-supervised learning for wearable-based momentary stress detection in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *7*(2), 1-23.