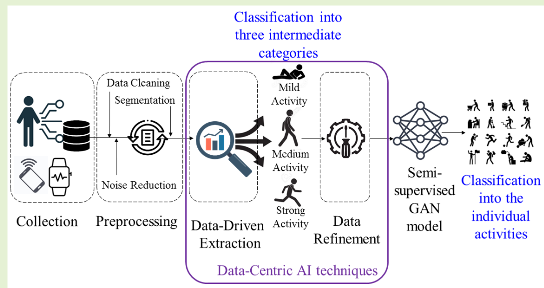
A Data-driven Feature Extraction Method Based on Data Supplement for Human Activity Recognition

Myung-Kyu Yi, *Member, IEEE,* and SEONG OUN HWANG, *Senior Member, IEEE*

***Abstract*—Human Activity Recognition (HAR) has garnered attention as a significant technology that can enhance the quality of human life. However, existing HAR works still face great challenges such as a shortage of labeled data and the difficulty of rebuilding a deep learning model whenever the application environment (e.g., user or sensor position) changes. To address these challenges, we propose a new data-centric approach for HAR by using a Semi-supervised Generative Adversarial Network (SGAN). To improve the accuracy of HAR, we propose a data supplement strategy that systematically improves data quality, rather than the model, by using data refinement and data-driven feature extraction techniques. The proposed HAR method applies simple SGAN to achieve considerably high accuracy with only a small fraction of the labeled data. Therefore, the proposed HAR method can reduce overhead from data labeling, which is a labor-intensive and time-consuming process for many HAR tasks. Moreover, the data-centric HAR method is robust even in scenarios when there is a change in person/sensor location. Experimental results show that our method improves accuracy by as much as 3% over state-of-the-art semi-supervised HAR methods with only 3% of the data being labeled, leading to comparable accuracy to state-of-the-art HAR methods based on supervised learning.**



***Index Terms*—human activity recognition, data-centric AI, wearable device, generative AI, semi-supervised learning**

I. Introduction

ADVANCEMENT in various technologies like wearable devices, the Internet of Things (IoT), and Artificial Intelligence (AI) enable many interesting applications. One notable example is Human Activity Recognition (HAR), which has recently emerged as one of the most popular research topics. The main purpose of HAR is to identify and predict human behavior from raw activity data collected by sensors. From the various sensor data or input signals, user behavior can be analyzed and utilized through data processing in various fields, such as healthcare, wellness, fitness, physical rehabilitation, detecting abnormal behavior, and fall detection [1]–[3]. Moreover, HAR is a critical component in the development of intelligent systems, enabling machines to understand and respond to human behaviors in real-time [4]–[7]. By integrating HAR algorithms, these robotic assistants can perceive and interpret human actions, allowing them to assist with tasks ranging from everyday household chores to providing support for elderly or disabled individuals, thereby enhancing their autonomy and quality of life [8]–[10].

Myung-Kyu Yi and Seong Oun Hwang are with the Department of Computer Engineering, Gachon University.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2021-0-00540)

Manuscript received May 19, 2023; revised August 26, 20XX.

Based on the type of data acquired, HAR approaches can be broadly classified into three main groups: sensor-based HAR, vision-based HAR, and pose-based HAR. By using wearable sensors such as accelerometers, and gyroscopes, a sensor-based HAR can identify users regular everyday activity or abnormal behavior, such as gait freezing [11]–[16], [18]. A vision-based HAR is a method for identifying and interpreting various human activities by extracting meaningful features or patterns form the visual data or information obtained from cameras or other vision sensors [19], [20]. One of the most prevalent vision-based HAR systems is pose-based HAR [21], [22]. Pose-based HAR focuses on identifying and comprehending human activities by examining and interpreting human poses that extract the spatial organization of the human body, such as skeletal tracking or keypoint identification. By integrating data from multiple modalities, including sensor-based (e.g., accelerometer, gyroscope), auditory (e.g., audio signals), and visual (e.g., video, pictures) data, multimodal HAR was proposed to enhance the reliability and accuracy of activity recognition systems [23]. Although various types of HAR are being developed, the most important and promising technology for supporting people in their daily lives is still sensor-based HAR considering the simplicity and convenience of human behavior monitoring. Therefore, our work focuses on the sensor-based HAR, which can be used to improve various aspects of healthcare, patient care, medical treatment,

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

TABLE I
COMPARISON BETWEEN MODEL-CENTRIC AI AND DATA-CENTRIC AI FOR HAR

Criterion	**Existing Works: MC-AI HAR**	**Our Approach: DC-AI HAR**
Fundamental Principle	Focuses on optimizing the performance of models	Focuses on optimizing the quality and quantity of data
Focus	Model architecture, model selection, parameter tuning	Data collection, data cleaning, data labeling, data augmentation
Strengths	Capable of solving complex problems using advanced models	Achieving good performance with simple models if the data is sufficient and of high quality, Easy to accommodate with wearable devices due to the simplified model
Weaknesses	Model performance can be limited in the absence of high-quality data, Complex models with large number of parameters, Fail to detect bias and under-representation, Challenges in applying wearable devices	Requiring additional data analysis and data cleaning processes

and ambient assisted living.

Earlier HAR methods were built on Machine Learning (ML) due to its accuracy and robustness across various HAR datasets. However, handcrafted methods are needed for ML techniques to extract features from datasets [18], [24], [25]. The development of these feature extraction methods requires domain knowledge, which can be a serious drawback. On the other hand, Deep Learning (DL) has gained more traction recently, compared to ML, because DL can achieve higher accuracy in general [11]–[16], [18].

The majority of existing HAR can be classified as Model-Centric AI based HAR (MC-AI HAR) [11]–[16]. In this approach, the emphasis is placed more on improving the HAR model rather than on the data. Achieving high accuracy requires complex HAR models with many parameters, which makes it difficult to apply HAR to IoT devices with limited memory and computational resources [17]. Although the model-centric HAR approach is useful in many circumstances, there are several challenges:

1) Model-centric HAR approaches do not pay ample attention to the comprehensive data exploration [18]. Data exploration is crucial for gaining insights, identifying patterns, and uncovering potential biases or anomalies within the data. As a result, MC-AI HAR approaches may fail to detect bias and under-representation, leading to performance degradation [18], [26].
2) A shortage of labeled data is one of the biggest challenges faced by model-centric HAR [26], [27]. Data labeling is an integral part of various ML and DL models in AI applications. However, it is also one of the most time-consuming and labor-intensive operations when we deploy an ML or DL model in the real world; it raises the overall operational costs of a HAR system. To solve this problem, semi-supervised learning based HAR method [28], [29] was proposed, but it failed to achieve the accuracy of supervised learning HAR methods [11], [13].
3) A HAR model trained in one scenario yields inaccurate results in another scenario because the performance of model-centric HAR methods can be affected by factors such as different users and sensor positions [30]–[32]. The domain-shift problem between a labeled source domain and an unlabeled destination domain was addressed by several transfer methods based on unsupervised learning [30]–[32]. However, the accuracy results of unsupervised methods are lower in most cases.

The emerging trend of Data-Centric AI (DC-AI) aims to address the limitations of traditional MC-AI approaches by enabling systems to learn from data instead of relying solely on complex AI models or algorithms [33], [34]. In this context, we introduce a novel approach for HAR, called the DC-AI based HAR (DC-AI HAR), leveraging the DC-AI methodology as shown in Table I. This method utilizes a simplified SGAN model to solve the shortage of labeled data while achieving high accuracy. To enhance the learning capabilities of the simplified SGAN model, we integrate advanced DC-AI techniques, including data-driven feature extraction and data refinement. This approach demonstrates the superiority of DC-AI methodology, especially in applications such as HAR, where the quality and sophistication of data play a critical role in model performance. Our contributions are summarized as follows.

1) To the best of our knowledge, this paper is the first to apply the DC-AI approach to HAR. Based on in-depth exploration and analysis of data [18], we focus on data refinement that systematically improves the quality of the data, rather than tuning the model. As a result, we can achieve an astonishing improvement of accuracy up to 9.9% compared to the baseline SGAN model. The proposed method presents an effective data engineering methodology that is generically applicable to many practical DL applications.
2) By applying the data-centric techniques to SGAN model, the proposed HAR method amazingly improves accuracy with 3% over the state-of-the-art SGAN method based on the MC-AI HAR [28]. As a result, the proposed HAR method can achieve an equivalent accuracy to those of supervised learning HAR methods (an F1 score of 0.002 higher than [41], and F1 scores of 0.054 and 0.039 lower than [13], [14], respectively). It should be noted that these accuracies were achieved with only 3% of the data being labeled which is quite a small fraction, greatly reducing overhead from the time-consuming and labor-intensive data labeling process for many HAR applications.
3) The proposed HAR method has an F1 score of 0.176 higher than that with the unsupervised learning HAR method [31], [32]. Moreover, the proposed SGAN model has 127× and 23× lesser discriminator and generator parameters compared to [28], guaranteeing higher

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

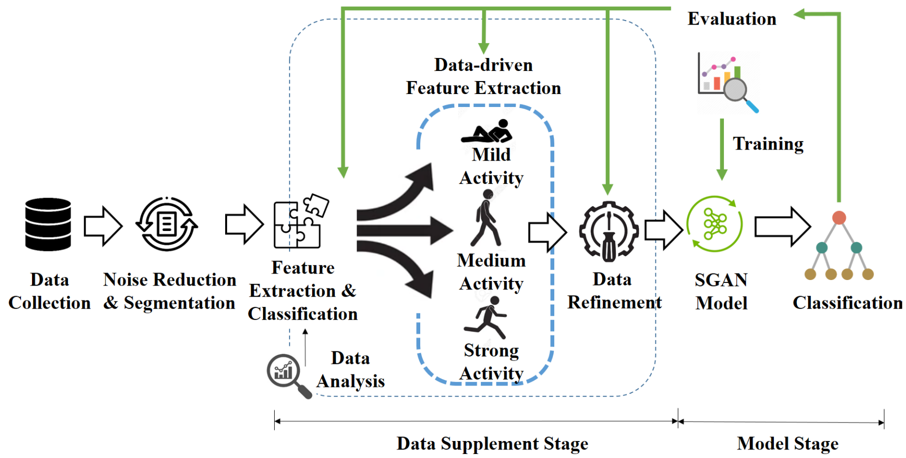


Fig. 1. Conceptual overview of the proposed data-centric HAR system

accuracy and less training time. Therefore, it can alleviate the domain-shift problem as well as solve the aforementioned problem of the shortage of labeled data, because retraining to adapt to the new environment can be carried out quickly with only 3% of the labeled data. For instance, when our proposed HAR method is used for a child instead of an elderly person, or if the wearable sensor changes from the left hand to the right hand, we can retrain the model with far less labeled data and still achieve acceptable recognition accuracy. This makes our proposed HAR method flexible and practical in real-life scenarios.

Our implementation is available in the public domain: https://github.com/kainos14/SGAN. The novelty of our work lies in the pioneering application of the DC-AI approach to HAR, a cutting-edge paradigm not previously explored in HAR studies. Existing HAR works rely heavily on predefined models, which might only capture some of the nuances of human activities, so they struggle to generalize well across different sensor types, environments, or user populations. We present a new DC-AI approach that can quickly adapt to changes in sensor configurations or evolving data distributions by combining a simple model and the proposed DC-AI techniques with small amounts of labeled data. This adaptability can lead to developing more resilient and dynamic recognition HAR systems ready to meet the ever-changing needs of real-world applications.

The rest of the paper is organized as follows. Section II reviews the recent state-of-art works on sensor-based HAR methods as the related works. Section III describes the details of the proposed data-centric HAR method. Section IV reports the experimental results and discussion on the achieved results. Section V concludes the paper.

II. Related Work

There have been various kinds of research on sensor-based HAR, as discussed in [11]–[16], [18], [35], which involve both ML and DL approaches. One of the notable HAR studies is the two-level HAR approach [15], [16] to reduce power consumption in order to improve HAR accuracy. Huang et al. [16] present a new HAR approach that includes a two-stage end-to-end CNN structure with data augmentation. Cho et al. [15] proposed a one-dimensional Convolutional Neural Network (CNN) for HAR that employs a divide and conquer-based classifier learning coupled with test data sharpening. These approaches are appropriate for sensor-based wearable HAR that requires high precision while consuming little power. But, two-level HAR approaches adds complexity to the overall system design and implementation. Moreover, these approaches often require more computational resources compared to single-level approaches. The ensemble and hybrid HAR techniques are the most often used classification methods. Several works [36]–[38] have presented efforts to achieve high-level accuracy in HAR by utilizing ensemble and hybrid approaches. Semwal et al. [36] proposed four hybrid DL models for the classification of gait activities, namely, CNN–Long Short-Term Memory-CNN (CNN-LSTM), CNN-Gated Recurrent Unit (CNN-GRU), LSTM–CNN, and LSTM–GRU, to provide the generic activity recognition framework and tune the performance. Thakur et al. [37] proposed a novel DL approach in which a convolutional autoencoders (AEs) with LSTM for HAR, namely, ConvAE-LSTM. Xia et al. [38] proposed a DL model that combines convolutional layers with LSTM. Neha Gaud et al. [39] proposed a hybrid deep learning-based model that is trained to recognize the various gestures. Three deep learning-based models, namely 1D CNN, CNN-LSTM, and CNN-GRU, are designed to test the various human mobility gestures. Manoj Kumar Sain et al. [40] proposed a novel deep learning architecture, Convolutional-Bidirectional LSTM, and GRU (ConvBiLSTM-GRU). Sravan Kumar Challa et al. [41] proposed the DL-based HAR model that combines CNN layers and BiLSTM units to simultaneously extract spatial and temporal sequence features from raw sensor data. The Rao-3 metaheuristic optimization algorithm has been adopted to identify the ideal hyperparameter values for the proposed DL model to enhance its recognition performance. Mohammed et al. [42] proposed a DL model called Multi-ResAtt (Multilevel Residual network with Attention). This model incorporates initial blocks and residual modules aligned

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

in parallel, so it can integrate a Recurrent Neural Network (RNN) with attention to extracting time-series features. Kosar et al. [43] proposed a novel hybrid DL model that combines the LSTM and the 2D CNN branches for the HAR system. This model runs parallel to receive the raw signals and their spectrograms to concatenate the features extracted at each branch and use them for activity recognition. Sakorn et al. [44] proposed the hybrid DL model with hybrid Squeeze-and-Excitation Residual blocks (SEResNet) combining a Bidirectional Gated Recurrent Unit (BiGRU) to extract deep spatio-temporal features hierarchically and to distinguish transitional activities efficiently. Ensemble and hybrid HAR techniques often require the training and inference of multiple models at the same time. This may significantly increase the computational requirements both in terms of memory and computing power. Due to the increasing computational load, real-time applications might be unable to achieve the appropriate performance.

Despite the significant improvements achieved by sensor-based HAR approaches, these supervised learning-based HAR systems require a considerable amount of accurately labeled training data. It can be expensive and labor-intensive to gather a comprehensive and accurately labeled dataset because HAR often is made up complex and diverse human behaviors. To overcome the disadvantages of supervised learning-based HAR, unsupervised learning-based HAR techniques have been proposed in [30]–[32]. Zhao et al. [32] proposed two new approaches for cross-domain HAR called local domain adaptation (LDA) and auto clustering LDA (ACLDA). Chen et al. [31] proposed a novel Stratified Transfer Learning (STL) framework for cross-domain HAR. Unsupervised learning-based HAR does not rely on labeled training data, which can be helpful in situations when labeled data is limited or expensive to obtain. However, unsupervised learning-based HAR approach has significant drawbacks compared to supervised learning-based HAR.

Firstly, the result from unsupervised learning methods is usually less accurate [30]–[32] compared to the supervised learning-based HAR methods. Secondly, it may have difficulty effectively identifying and distinguishing between human activities, especially if the data is complicated or the activity groups are similar [45]. Finally, noisy or untrustworthy sensor data might result in erroneous or ambiguous activity clustering, resulting in poorer accuracy [46]. To overcome the drawback of supervised and unsupervised learning-based HAR methods, Zilelioglu et al. [28] presented an SGAN framework for HAR applications that includes temporal convolutions. For the evaluation of the proposed HAR methods, the public datasets called the PAMAP2 [47], Opportunity-locomotion, and LISSI HAR were used. The experimental results showed that the proposed HAR method can achieve high classification performance by reaching improvements of up to 25% even with a 3% amount of labeled data. Another SGAN framework is given in [29] for cross-subject domain transfer semi-supervised domain adaption. However, experimental results showed significantly lower classification performance compared with previous works.

However, the proposed HAR method still relies on model architecture so it has to depend on domain-specific models. It is necessary to gather new data and retrain the model to adapt them to new domains, which can be time-consuming and expensive [33], [34]. Moreover, the convolutional layers in the SGAN model can be computationally expensive and require a significant amount of memory for storing the parameters [48]. To overcome the drawback of model-centric HAR which heavily depends on the AI model, we propose a semi-supervised learning-based HAR by applying the DC-AI technique which improves the accuracy results through data enhancement techniques.

III. The Proposed Data-Centric HAR Method

This section provides overviews of the proposed data-centric HAR method, the data-driven feature extraction, data refinement, and steps in the SGAN model.

A. System Overview

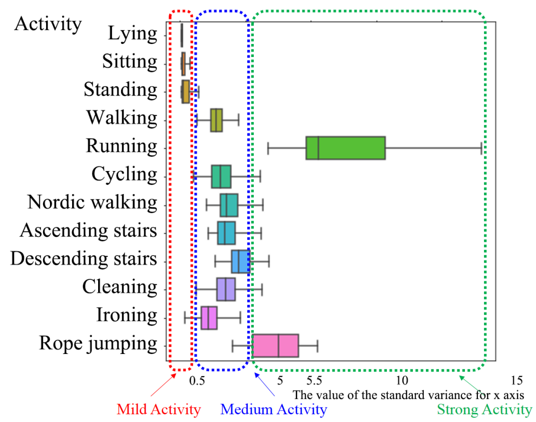


Fig. 2. Distribution of activities based on standard variances on the x-axis of the PAMAP2 dataset ($T_L = 0.5, T_H = 5.5$)

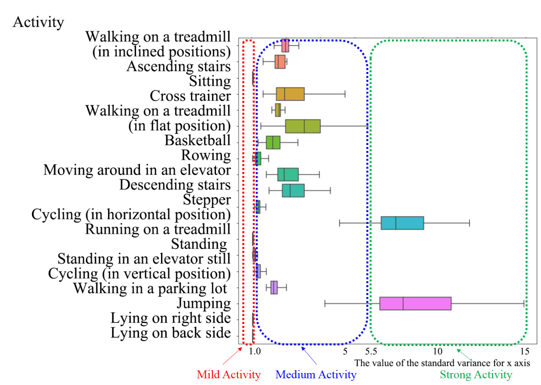


Fig. 3. Distribution of activities based on standard variances on the x-axis of the DSADS dataset ($T_L = 1.0, T_H = 5.5$)

In Fig. 1, raw data are first collected from wearable sensors and then preprocessed to reduce noise [49], [50]. The preprocessed data are divided into segments using a sliding window with an overlapping region to minimize the impact of noise and to reduce processing costs for ML or DL algorithms. To obtain high recognition accuracy, feature extraction is carried out to obtain statistical feature vectors such as the mean and

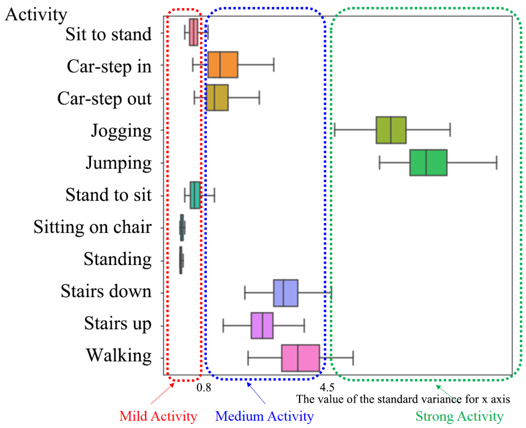


Fig. 4. Distribution of activities based on standard variances on the x-axis of the ModiAct dataset ($\boldsymbol{T_L = 0.8, T_H = 4.5}$)

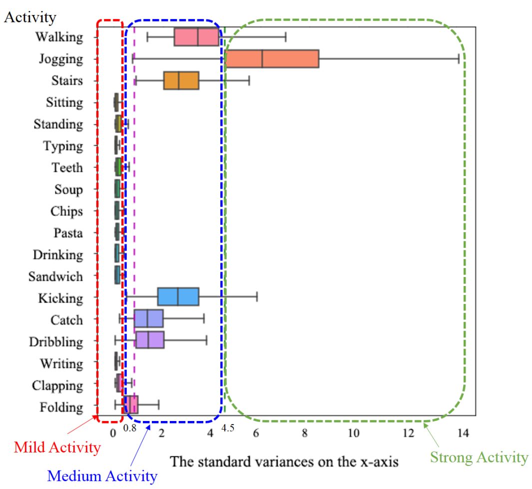


Fig. 5. Distribution of activities based on standard variances on the x-axis of the WISDM dataset ($\boldsymbol{T_L = 0.8, T_H = 4.5}$)

the standard variance for each windowed set of data after segmentation. Note that this stage is commonly present in existing HAR systems [18]. Finally, these feature vectors via the data-driven feature extraction process are fed into the appropriate classifier.

B. Data-driven Feature Extraction

The novelty of our work lies in the detailed classification of human behaviors into mild, medium, and strong categories, moving beyond the traditional dynamic/static classification [15]–[18] through an exploration of data characteristics [18]. This segmentation simplifies data classification complexity. Therefore, the proposed data-driven feature extraction technique can improve HAR accuracy by employing a divide-and-conquer strategy [15], [16]. In general, there are three primary human activity categories according to the movement intensity: mild (standing, sitting, lying down), medium (walking, descending stairs, ascending stairs, etc.), and strong (e.g., jumping, running). Based on this, we propose classifying the incoming sensor data into three primary activity categories by utilizing standard variances on the x-axis, σ_x, as shown in Eq. (1):

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad (1)$$

Figures 2, 3, 4, and 5 show standard variances on the x-axis for accelerometer data from the PAMAP2 [47], the DSADS [24], the MobiAct [51], and WISDM [52] datasets, respectively. These datasets were used in many previous HAR papers [11]–[14], [25], [28], [31], [32]. We can see that all activities are classified as mild (red dotted rectangle), medium (blue dotted rectangle), or strong (green dotted rectangle). Based on these observations, we build a two-stage activity recognition procedure where primary activity categories (mild, medium, and strong) are recognized first, and individual activities are recognized using the SGAN. Note that we do not directly recognize individual activities using a single classifier because that would introduce classification overhead and lower the accuracy [15], [16]. With this approach, we can also reduce the number of classes to be identified by each SGAN model, effectively reducing the training and inference overhead. It is worth noting that the sensor data can also be categorized into three principal activity groups using the y-axis and z-axis distributions, as they show behavior similar to the x-axis data. As a result, the proposed method can be utilized no matter how the position of the sensors. This shows that the proposed method can be used regardless of the position of the sensors [17]. Algorithm 1 shows an overview of the proposed data-driven feature extraction technique. Let T_L and T_H denote the threshold values to classify into the primary activity categories, respectively.

Algorithm 1 : Data-Driven Feature Extraction

Input: Raw sensor data D
Output: Classified dataset $X_{mild}, X_{medium}, and\ X_{strong}$
Initialization
Raw sensor data are segmented into fragments D_i
$D = D_1, D_2, ..., D_N$
Get the value of the standard variance $\sigma_x[D_i]$ for each D_i.
Loop
for $(i = 1$ to $N)$ **do**
  if $\sigma_x[D_i] < T_L$ **then**
    Classified into X_{mild} as the mild activity
  else if $\sigma_x[D_i] > T_H$ **then**
    Classified into X_{strong} as the strong activity
  else
    Classified into X_{medium} the medium activity
  end if
end for

Raw sensor data can be segmented into fragments, then obtain the statistical feature vectors such as the values of mean and the standard variance for each fragment. If the standard variance on the x-axis, σ_x, of the data fragment is less than T_L, it will be classified into mild activity. If the standard variances on the x-axis, σ_x, of the data fragment is larger than T_H, it will be classified into strong activity. Others will be classified into medium activity. The values of T_L and T_H can be easily obtained from a dataset by exploratory data analysis method [18] by Tensorflow matplotlib [56].

C. Proposed Data Refinement Technique

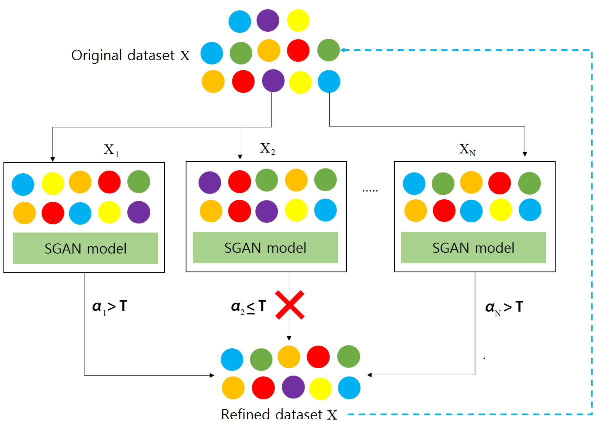


Fig. 6. Schematic of the proposed data refinement technique

Algorithm 2 : Data Refinement

Input: Original dataset X

Output: Refined dataset $X^{'}$

Let X be the dataset and let y be the corresponding labels
N : The number of groups

Initialization

Generate N datasets of the same size as the original by random sampling, $(X = X_1, X_2, .., X_N)$

Each dataset : $X_i = x_1, x_2, ..., x_n$ and $y_i = y_1, y_2, ..., y_n$

Loop

for $i = 1$ to N **do**

Train the SGAN model on X_i

Get probability $p_i = \sum(y_i - \hat{y}_i)$

where y_i is the actual label and $\hat{y}_i$ is the predicted label

Calculate accuracy $a_i = \frac{1-p_i}{n}$ for X_i

end for

Get average accuracy T from $\frac{\sum_{i=0}^{N} a_i}{N}$

if $(a_i > T)$ **then**

Include X_i in $X^{'}$

end if

Imbalanced data refers to datasets that have an uneven distribution of target classes. This is a critical problem because it can negatively affect the accuracy of the DL model, degrading HAR classification [26]. To reduce issues caused by data imbalance, we propose a data refinement technique that adopts a strategy of balancing labeled data based on class and selecting a good sample for the SGAN model. Algorithm 2 and Fig. 6 show an overview of the proposed data refinement technique. In the proposed technique, the SGAN model is used for both data refinement and HAR training/testing. We denote X_i and α_i as the ith dataset and the accuracy value of X_i, respectively. The key idea of data refinement is to produce a more stable and accurate sample through random sampling. The use of poor quality data can lead to severe degradation in the results. To avoid this problem, we generate multiple new datasets N of the same size as the original by random sampling (i.e., $X_1, X_2, .., X_N$). Each dataset has the same number of labels; we train the SGAN on X_i, then get the value of α_i. From the values of α_i, we get the value of average accuracy, T, for all datasets. Finally, only datasets that are higher than T are selected as the SGAN model input. The process can be iterated several times to select datasets with good quality. Since there are many ways to solve the class imbalance problem through resampling methods, including over-sampling and under-sampling [26], performance in terms of accuracy from the SGAN model can be further improved if a more advanced sampling method is applied in the proposed HAR method.

D. Semi-Supervised GAN

Fig. 7 shows the architecture of semi-supervised GAN. We followed the guidelines to build the baseline SGAN model extended from the vanilla GAN [53], [54]. Two completely connected networks are utilized for the generator and discriminator in the original improved GAN [54], but we adjust the number of dense layers to improve the accuracy performance of the HAR system. For simplicity, the architecture of the proposed SGAN model consists of the generator and discriminator with only simple dense layers because we focus on data refinement rather than model tuning. Therefore, the training time is much shorter than the GAN model that utilizes convolutional layers or Long Short-Term Memory (LSTM) layers. We only adjust the number of dense layers in the SGAN model. In the proposed SGAN model, a supervised discriminator, an unsupervised discriminator, and the generator are all simultaneously trained as part of the SGAN model, which is an extension of GAN architecture. The reason for selecting the SGAN is that it only uses a small amount of labeled data to achieve good accuracy. This combines the advantages of supervised and unsupervised learning, and therefore, a huge amount of labeled data is not required. Furthermore, the SGAN model is suitable for the proposed data-centric HAR method, especially in practical scenarios where data labeling is difficult. In the SGAN [54], discriminator D is changed into a classifier with $K+1$ classes. The numbers of activities are denoted by the first K class, and the $K+1th$ class is a fake sample from the generator G.

$$L_{\mathrm{s}} = -\mathbb{E}_{x,y\sim p_{\mathrm{data}}(x,y)} \log p_{\mathrm{model}}\left(y|x, y < K+1\right) \quad (2)$$

$$L_u = -\Big\{ \mathbb{E}_{x\sim p_{\mathrm{data}}(x)} \log[1 - p_{\mathrm{model}}\left(y = K+1|x\right)] \quad (3)$$
$$+ \mathbb{E}_{x\sim G(z)} \log[p_{\mathrm{model}}\left(y = K+1|x\right)] \Big\}$$

In equations (2) and (3), the supervised loss function, L_s, is the cross-entropy loss function as in a supervised learning setting with K classes, while the unsupervised loss function, L_u, is the usual GAN loss function. In optimization, the total loss, L_SGAN, is used. This architecture can reduce the amount of labeling required to achieve a certain level of

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

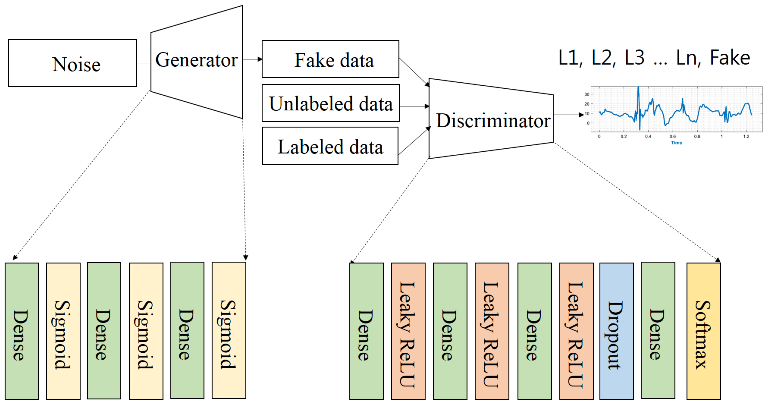


Fig. 7. The baseline SGAN model extended from the vanilla GAN

TABLE II
DEFINITIONS OF DIFFERENT LEVEL OF ACTIVITY FOR WEARABLE DATASETS

Dataset	Activity	# subjects	# labels	Class labels in Activity
PAMAP2	Mild	9	3	Lying, Sitting, Standing
	Medium		7	Walking, Cycling, Nordic walking, Ascending stairs, Descending stairs, Vacuum cleaning, ironing
	Strong		2	Running , Rope jumping
DSADS	Mild	8	4	Sitting, Standing, Lying on back and on right side
	Medium		13	Ascending and descending stairs, Standing in an elevator still, Moving around in an elevator, Walking in a parking lot, Walking on a treadmill with a speed of 4 km/h (in flat and 15 deg inclined position), Exercising on a stepper, Exercising on a cross trainer, Playing basketball, rowing, Cycling on an exercise bike in horizontal and vertical position
	Strong		2	Running on a treadmill with a speed of 8 km/h, Jumping
MobiAct	Mild	19	4	Sit to stand, Stand to sit, Sitting on chair, Standing
	Medium		5	Car step in, Car step out, Walking, Stairs up, Stairs down
	Strong		2	Jogging, Jumping

accuracy by allowing unlabeled real data to participate in the learning. From (2) and (3), the cost function to train classifier L_{SGAN} is as follows [55]:

$$L_{SGAN} = L_s + L_u \tag{4}$$

Data are represented by x in the previous notations, whereas labels are represented by y, and $p_{\text{model}}\,(y = K + 1|x)$ denotes the probability that x is fake data.

IV. Experimental Results and Analysis

A. Evaluation Metric and Dataset

This section describes the experimental results and analysis for the proposed DC-AI HAR method. The evaluation metrics used in benchmarking our method are given in Eqs. (5)-(8).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

(TP : True Positive, TN : True Negative, FP : False Positive, FN : False Negative)

We chose the PAMAP2 [47], the DSADS [24], and the MobiAct [51] datasets because they are openly available and have been used in numerous previous articles on the topic of HAR [11]–[14], [25], [28], [31], [32]. The PAMAP2 dataset contains data from 18 different physical activities performed by nine people while wearing three accelerometers and a heart rate monitor. The accelerometers were placed on the wrist, chest, and ankle (3 Colibri wireless inertial measurement unit with sampling frequency 100Hz). Three male and one female subjects (ages 25 to 30) participated in this data collection. Each individual followed a prescribed protocol and completed 12 different activities; some completed six extra tasks. The DSADS dataset includes information from 19 distinct physical activities carried out by eight subjects (four female, four male, between the ages 20 and 30) wearing body sensors. The tri-axial acceleration, gyroscope, and magnetometer data were captured using 5 MTx orientation trackers placed on the torso, right arm, left arm, right leg, and left leg. Sensor units are calibrated to acquire data at 25 Hz sampling frequency. Each subject's total signal time for each exercise was five minutes. The MobiAct dataset contains raw sensor data from a participant's smartphone collected during various falls and everyday life activities. The participants chose which pocket to put the phone in, and it was placed randomly within the pocket.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

The dataset was gathered from 19 participants participating in 11 distinct activities. Data were recorded from the accelerometer, gyroscope, and orientation sensors of a Samsung Galaxy S3 smartphone with the LSM330DLC inertial module (3D accelerometer and gyroscope). For generating the MobiAct dataset, 57 subjects (42 men and 15 women) were recorded. The subjects' ages ranged between 20 and 47 years, the height ranged from 160 cm to 189 cm, and the weight varied from 50 kg to 120 kg.

To achieve a fair comparison, we applied five-fold cross-validation for all datasets and used the same number of classes for each dataset as same to [11]–[14], [25], [28], [31], [32], [35], [42]–[44] as shown in Table II. The PAMAP2, the DSADS, and the MobAct datasets were repeatedly divided into training and validation sets for cross-validation. All of the experiments were conducted on a machine with an Intel Core i7-6700 CPU and 8GB RAM. We made considerable use of the TensorFlow library [56] for building an existing ML model and our proposed HAR method.

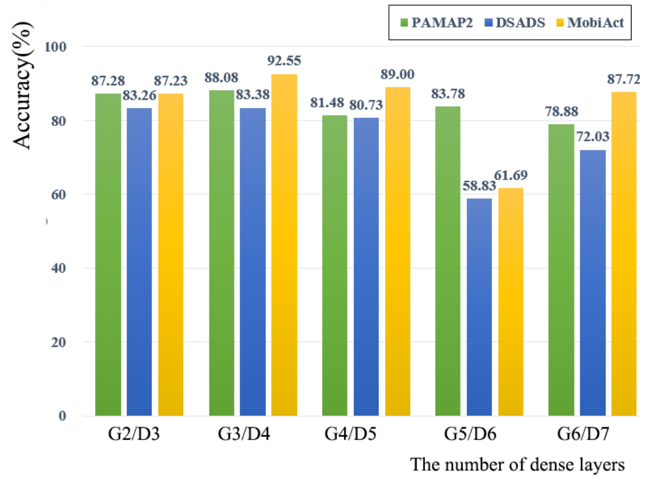


Fig. 8. Performance of the SGAN model when increasing the number of dense layers

B. Experimental Results

To evaluate the proposed DC-AI HAR method, we conducted an ablation study. Firstly, we measure the performance of the baseline SGAN model without applying the DC-AI techniques. Then, we investigate the impact of the number of dense layers and neurons on the SGAN model with varying amounts of labeled data. Finally, we examine the accuracy performance of the SGAN model combined with the proposed data refinement and data-driven feature extraction techniques.

TABLE III
CLASSIFICATION ACCURACY FOR THREE PRIMARY ACTIVITY CATEGORIES BASED ON DATA-DRIVEN FEATURE EXTRACTION AND VARIOUS ML CLASSIFIERS

Method	PAMAP2 Accuracy	DSADS Accuracy	MobiAct Accuracy	Training
Super Vector Machine	95.8%	99.9%	91.2%	YES
Decision Tree	98.8%	99.8%	98.3%	YES
Random Forest	99.6%	99.9%	98.8%	YES
Logistic Regression	99.3%	99.9%	98.8%	YES
Data-driven Feature Extraction	95.7%	99.9%	95.1%	NO

1) Performance Evaluation for Data-driven Feature Extraction: Table III shows the accuracy performance of the data-driven feature extraction (Section III.B) and various ML classifiers for three primary activity categories (mild, medium, and strong activities) on the PAMAP2, the DSADS, and the MobiAct datasets, respectively. Recognition accuracy from the proposed data-driven feature extraction technique is slightly lower compared to other ML classifiers using the PAMAP2 and MobiAct datasets but shows good accuracy above 95%. It also achieves the same accuracy as other ML classifiers on the DSADS dataset. Note that data-driven feature extraction only requires a very small amount of computation, and does not need any training. It can be applied to a HAR system immediately and increases accuracy by reducing the number of categories to be classified by the SGAN model.

2) Impact of the Number of Dense Layers on the SGAN model: Fig. 8 shows the performance of the SGAN model when increasing the number of dense layers with 3% of the data being labeled. We denote a generator with n dense layers and a discriminator with m dense layers as Gn and Dm, respectively. For example, $G2/D3$ indicates a generator with two dense layers and a discriminator with three dense layers. When increasing the number of dense layers, the accuracy of the SGAN model increases, as shown in Fig. 8. We obtained the best accuracy (88.08%) with $G3$ and $D4$ while classifying the PAMAP2 dataset. Similarly, the best accuracies were 83.38% with $G3$ and $D4$, and 92.55% with $G3$ and $D4$ while using the DSADS and MobiAct dataset, respectively. We can see that more features be extracted from the SGAN model as the number of dense layers increases. However, we can observe that increasing of dense layer actually caused the reduction in accuracy of HAR after $G3/G4$.

Although the SGAN model produced outstanding results on a wide range of semi-supervised learning tasks, training them is extremely unstable, challenging, and hyperparameter sensitive. When a generator can only produce a single type (or a short-range) of output, model collapse occurs [57]. Mode collapse is a common problem in GANs, where the generator produces similar or identical samples instead of diverse and realistic ones. Firstly, this can happen when the discriminator becomes too powerful and rejects most of the generator's outputs or when the generator finds a few samples that can fool the discriminator into following them. Secondly, increasing the number of layers in a GAN can lead to the problem of gradient vanishing or exploding. Finally, deeper GANs can be prone to overfitting the training data, causing the model to focus too much on specific modes and lack diversity. Therefore, it becomes more challenging for the generator and discriminator to update appropriately, which can result in mode collapse. So, we can see that model collapse cannot be avoided when the number of dense layers on the SGAN model increases beyond $G3/G4$.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

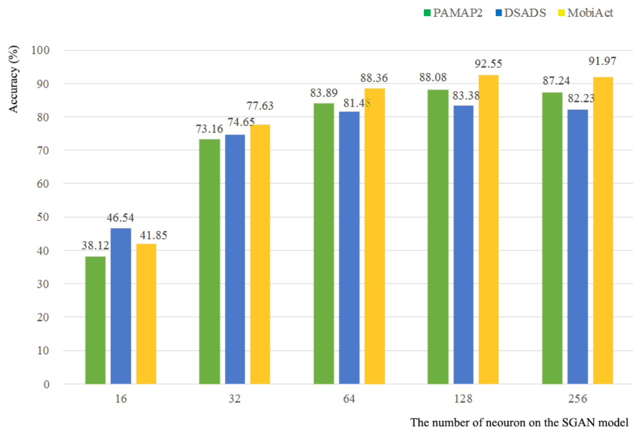


Fig. 9. Performance of the SGAN model when increasing the number of neurons

TABLE IV
PERFORMANCE OF THE BASELINE SGAN MODEL FROM INCREASING THE AMOUNT OF LABELED DATA

Labeled Data	PAMAP2	DSADS	MobiAct
1%	87.41%	64.49%	86.03%
2%	87.73%	70.48%	88.87%
3%	88.08%	83.38%	92.55%
4%	89.57%	85.91%	93.96%
5%	90.65%	86.23%	94.14%
6%	90.98%	89.69%	94.96%
7%	91.57%	90.40%	95.11%
8%	92.22%	91.38%	95.53%
9%	92.69%	91.89%	96.22%
10%	93.25%	92.48%	97.36%

3) Impact of the Number of Neurons on the SGAN model: Fig. 9 shows the performance of the SGAN model when increasing the number of neurons on the baseline SGAN model with 3% of the data being labeled. When the number of neurons increases, the SGAN model's accuracy increases. We obtained the best accuracy (88.08%) with 128 neurons while classifying the PAMAP2 dataset. Similarly, the best accuracies were 83.38% with 128 neurons and 92.55% with 128 neurons on the DSADS and MobiAct datasets, respectively. We can see that more features are extracted from the SGAN model as the number of neurons increases. When the number of neurons on the SGAN model exceeds 128, model collapse cannot be avoided, similar to Fig. 8.

4) Impact of the Amount of Labeled Data: Table IV shows the performance of the baseline SGAN model when increasing the percentage of labeled data. We can see that the accuracy of the SGAN increases as the amount of labeled data increases, which happened with all datasets. Through this experiment, we obtained almost 90% accuracy with only 3% to 7% of the data labeled for the baseline SGAN model. Fig. 10, 11, and 12 show the improved performance of the baseline SGAN model when integrated with the proposed DC-AI HAR techniques. We observe a consistent accuracy increase as more labeled data is utilized, which aligns with the observation in Table IV. For the PAMAP2 dataset, we observed a notable improvement in the accuracy performance of the baseline SGAN models after applying the proposed data refinement technique, as shown in

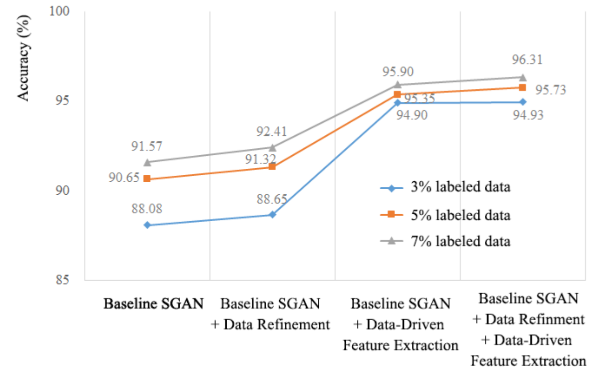


Fig. 10. Accuracy from the proposed DC-AI HAR method on the PAMAP2 dataset

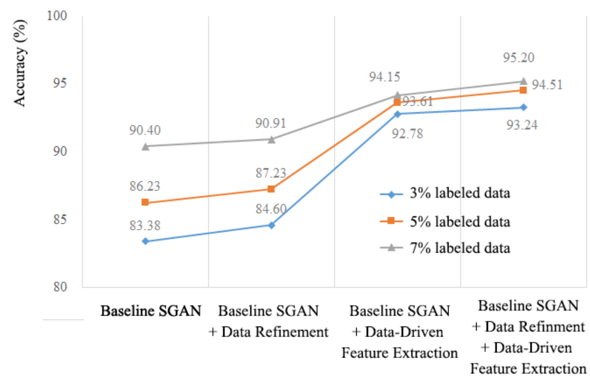


Fig. 11. Accuracy from the proposed DC-AI HAR method on the DSADS dataset

Fig. 10. Specifically, the accuracy of the baseline SGAN model increased by 0.57%, 0.67%, and 1.14% for 3%, 5%, and 7% labeled data, respectively. Additionally, we can observe that the proposed data-driven feature extraction technique boosted accuracy by 6.82%, 4.70%, and 4.63% for 3%, 5%, and 7% labeled data, respectively. Finally, the proposed DC-AI HAR method on the PAMAP2 dataset achieved accuracies of 94.93%, 95.73%, and 96.31% with 3%, 5%, and 7% labeled data, respectively. As shown in Fig. 11, the accuracy of the baseline SGAN model on the DSADS dataset improved after applying the proposed data refinement technique, resulting in gains of 1.22%, 1.00%, and 0.51% for 3%, 5%, and 7% labeled data, respectively. The proposed data-driven feature extraction technique increased accuracy by 9.40%, 7.38%, and 3.75% for 3%, 5%, and 7% labeled data, respectively. The proposed DC-AI HAR method on the DSADS dataset achieved accuracies of 93.24%, 94.51%, and 95.20% for 3%, 5%, and 7% labeled data, respectively. For the MobiAct dataset, the accuracy of the baseline SGAN model saw improvements of 0.60%, 0.30%, and 0.25% after applying data refinement, as shown in Fig. 12. We can see that the proposed data-driven feature extraction technique enhanced the accuracy by 2.53%, 1.90%, and 1.90% for 3%, 5%, and 7% labeled data, respectively. The proposed DC-AI HAR method achieved accuracy rates of 95.74%, 96.03%, and 96.34% on the MobiAct dataset for 3%, 5%, and

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

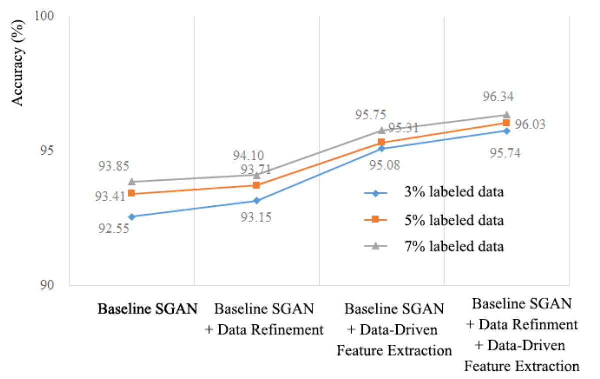


Fig. 12. Accuracy from the proposed DC-AI HAR method on the MobiAct dataset

TABLE V
F1 SCORES OF THE SGAN CLASSIFICATION IN THE CROSS-SUBJECT SETTING

Dataset	Subject	[17]	This Work	Labeled Data
PAMAP2	S1	0.912	0.929	13.10%
	S2	-	0.934	15.20%
	S3	-	0.857	8.80%
	S4	-	0.874	11.47%
	S5	-	0.913	13.05%
	S6	-	0.887	12.60%
	S7	-	0.901	10.92%
	S8	0.843	0.869	14.20%
	Mean	0.877	**0.895**	12.46%
DSADS	S1	-	0.900	12.5%
	S2	-	0.898	12.5%
	S3	-	0.928	12.5%
	S4	-	0.912	12.5%
	S5	-	0.925	12.5%
	S6	-	0.885	12.5%
	S7	-	0.908	12.5%
	S8	-	0.905	12.5%
	Mean	-	**0.908**	12.5%
MobiAct	S2	-	0.915	3%
	S3	-	0.924	3%
	Mean	-	**0.919**	3%

7% labeled data, respectively. These results demonstrate that the proposed DC-AI HAR method significantly enhances the accuracy of the baseline SGAN model through data refinement and data-driven feature extraction.

5) Impact of Cross-Subject Training: We investigate the impact of utilizing the labeled and unlabeled data from different people on the performance of the proposed DC-AI HAR method. The proposed DC-HAR method is trained in a cross-subject setting utilizing labeled data from a single subject. The training data from the other subjects are used as unlabeled data. For the MobiAct dataset, only two participants were selected because there were too many participants to include all of them. Table V shows the F1 scores of the SGAN classification in the cross-subject setting on the PAMAP2, DSADS, and MobiAct datasets. We can see that the proposed DC-AI HAR method achieved F1 scores of 0.895, 0.908, and 0.919 for the PAMAP2, DSADS, and MobiAct datasets, respectively. In particular, the proposed DC-AI HAR method shows improved F1 scores of 0.017 and 0.018 compared to the SGAN model used in [28] for S1 and S8 participants, respectively. Participants in an experiment may perform the same activities differently due to variations in body size, movement patterns, and personal habits. Also, they have different habits and preferences when performing activities. For example, some individuals may walk faster or slower, affecting the sensor data's features. It can make it challenging for a model trained on one subject's data to recognize other subjects' activities accurately [30]–[32].

6) Comparison with Results of Previous Studies: Table VI shows the overall performance of the proposed data-centric HAR method to numerous previously reported results that used supervised and unsupervised HAR techniques. The existing HAR method using the DL model based on a supervised technique showed the best F1 score of 0.949 [41], 0.987 [12], and 0.996 [25] on the PAMAP2, the DSADS, and the MobiAct datasets, respectively. Compared to our method, supervised HAR techniques have slightly lower or higher accuracy [11], [12], [25]. But in practical scenarios, the user of the HAR system may change from use by children to use by elderly people, or the placement of the HAR sensor can shift from the right hand to the left. The supervised learning method requires labeling of all data before training begins in order to achieve high accuracy, and the DL model has to be rebuilt whenever the HAR scenario changes [27]. This creates many obstacles to developing and upgrading a HAR system due to time-consuming data labeling. Our approach only needs 3% of the data to be labeled to achieve similar performance, with only 0.054, and 0.039 F1 scores drop on the DSADS, and the MobiAct datasets, respectively. Moreover, we can see that the proposed HAR method with only 3% labeled data improves the 0.008 F1 score on the PAMAP2 dataset.

On the other hand, an unsupervised HAR method [31], [32] does not require labeled data. Zhao et al. [32] and Chen et al. [31] presented classification results using an unsupervised HAR method for cross-person and cross-position HAR tasks. In a cross-person HAR task, best F1 score from the unsupervised HAR method was 0.731 and 0.757 on the PAMAP2 and the DSADS datasets, respectively. In a cross-position HAR task, the best F1 score of the unsupervised HAR method was 0.440 and 0.667 on the PAMAP2 dataset and the DSADS dataset, respectively [31], [32]. Note that this approach yielded poor accuracy in HAR applications compared to our method. Based on the above results, we can see that the proposed HAR method is both practical and flexible because it can be applied right away on any new dataset after a small amount of data labeling (3%) and training.

Table VII shows the number of parameters of the SGAN model used in [28] and our proposed HAR method on the PMAP2 dataset, respectively. Our proposed SGAN model has 127× and 23× lesser discriminator and generator parameters compared to [28]. The accuracy of the DL model can be improved when a convolutional layer and a dense layer are used together [28], because more features can be extracted from the dataset. However, the number of parameters in this DL model also increases drastically. The increase in parameters adds to the complexity of the DL model and requires more data and computational resources for training. Moreover, the inference also consumes more memory and takes longer time, which

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

TABLE VI
F1 SCORES OF THE DIFFERENT SGAN METHODS

	Types of Learning	PAMAP2	DSADS	MobiAct	Model
[11]	Supervised	0.943	–		CNN
[41]	Supervised	0.949	–		CNN + BiLSTM
[25]	Supervised	–	–	0.996	Ensemble
[44]	Supervised	–	–	0.989	SEResNet-BiGRU
[42]	Supervised	0.929	–	–	BiGRU
[35]	Supervised	0.922	–	–	CNN
[12]	Supervised	–	0.987	–	CNN
[43]	Supervised	–	0.929	–	2D CNN-LSTM
[13]	Supervised	0.937	–	–	CNN
[14]	Supervised	0.914	–	–	CNN
[32]	Unsupervised (Cross-person) Unsupervised (Cross-position)	0.731 0.426	0.757 0.667	–	ACDAL
[31]	Unsupervised (Cross-position)	0.440	0.390	–	STL-SAT
[28]	Semi-supervised	0.921	–	–	SGAN with 3% labeled data
[29]	Semi-supervised	0.830	–	-	SGAN with 11% labeled data
This work	Semi-supervised	**0.951**	**0.933**	**0.957**	SGAN with 3% labeled data

may result in restrictions on HAR system development and reduce the HAR system's utilization [58].

The deployment of HAR systems on wearable devices offers numerous advantages, including the convenience of wearability, the precision of gathered data, the ability to collect data in real-time, and the capacity for emergency response. However, implementing these systems on wearables faces significant challenges due to such devices' limited memory and computational resources. For instance, wearable devices like the STM32 Nucleo-144 are equipped with a mere 1 MB of flash memory, necessitating the development of highly efficient deep learning models that maintain high accuracy with fewer parameters [17]. Our proposed DC-AI HAR method used a much smaller number of parameters (resulting in a memory size of only 1039.68 KB) and was able to achieve 0.03 higher F1 scores compared to [28] as shown in Table VI. This shows that the proposed HAR method is lightweight (using much fewer parameters) and has good accuracy. This makes it very suitable to be implemented on constrained devices like wearable devices or microcontrollers to carry out the desired HAR tasks.

C. Discussion

This section discusses the potential application of the proposed DC-AI HAR approach to vision-based HAR systems. The proposed HAR method uses data refinement, data-driven feature extraction, and a simple SGAN to achieve high accuracy with limited labeled data. The proposed data-driven extraction techniques are effective for sensor-based wearable data but pose challenges for human action recognition in video and radar data due to the complexity of such data. Video data has many pixels per frame, and radar data contains intricate signal patterns over time, making statistical analysis difficult. Moreover, sensors may collect these data types under various environmental conditions such as lighting, weather, camera, or radar settings, which can affect the data. Finally, these data types may represent the same action differently, even for identical actions. Dealing with this variability may require complex statistical models, but they can be simplified using conventional ML techniques to classify activities as mild, medium, or strong. The proposed data refinement techniques and SGAN model can be applied to the vision-based HAR systems regardless of the data types. We plan to extend our DC-AI HAR work to a video-based HAR system by extracting meaningful patterns from cameras or other vision sensors to enhance HAR accuracy.

TABLE VII
NUMBER OF PARAMETERS OF SGAN MODELS ON THE PAMAP2 DATASET

Study	Networks	# Parameters
[28]	Discriminator	14,580,172
	Geneator	3,453,444
This work	Discriminator	114,828
	Generator	151,716

V. CONCLUSION

Although HAR has emerged as a very interesting application with a big advance in AI, it has suffered the inherent problems of data shortage and domain drift. To address these issues, we proposed a new SGAN-based HAR method using a DC-AI approach, which is quite different from existing HAR methods that rely on heavy model fine-tuning. The proposed HAR method employs a data supplement strategy to significantly improve data quality, greatly reducing labor-intensive tasks such as data labeling and model rebuilding. The proposed data-driven HAR method improves accuracy by as much as 3% over state-of-the-art semi-supervised HAR methods [28] with only 3% of the data being labeled, leading to comparable accuracy to state-of-the-art HAR methods based on supervised learning as well as dealing well with the data-shortage problem. Also, the SGAN model used in the proposed HAR method has a better performance, and 127x and 23x fewer discriminator and generator parameters compared to [28]. This shows that our approach has a better performance and is more lightweight compared to [28], showing that our approach efficiently alleviates the domain-drift problem.

The proposed DC-AI HAR approach has outstanding results by using simple models, but mainly by focusing on recognizing basic actions such as walking, running, and sitting. In the

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

future, we plan to extend our work to identify more complex activities closer to real-life scenarios, such as cooking, swimming, exercising, and daily tasks.

REFERENCES

[1] N. Gaud, M. Rathore, and U. Suman, "Human Gait Analysis and Activity Recognition: A Review," In Proc. of the 2023 IEEE Guwahati Subsection Conference (GCON), Guwahati, India, pp. 1-6 , 2023.

[2] Jain, Rahul , and Vijay Bhaskar Semwal, "A novel feature extraction method for preimpact fall detection system using deep learning and wearable sensors,"*IEEE Sensors Journal*, Vol. 22(23), pp. 22943-22951, 2022

[3] Vijay Bhaskar Semwal, Abhishek Kumar, Pankaj Nargesh, Vaibhav Soni, "Tracking of Fall Detection Using IMU Sensor: An IoHT Application." Machine Learning, Image Processing, Network Security and Data Sciences, Springer Nature Singapore, 10-11 December 2023

[4] Vijay Bhaskar Semwal, Youngshik Kim, Vishwanath Bijalwan, Astha Verma, Ghanapriya Singh, Neha Gaud, Hangyeol Baek, and Abdul Manan Khan, "Development of the LSTM Model and Universal Polynomial Equation for all the Sub-phases of Human Gait", IEEE Sensors Journal, Vol. 23(14), pp. 15892 - 15900, 2023.

[5] Vijay Bhaskar Semwal, Neha Gaud, Praveen Lalwani, Vishwanath Bijalwan, and Abhay Kumar Alok, "Pattern identification of different human joints for different human walking styles using inertial measurement unit (IMU) sensor", Artificial Intelligence Review, Vol.55(2), pp. 1149-1169, 2021

[6] Vijay Bhaskar Semwal, Rahul Jain, Pushkar Maheshwari, and Saksham Khatwani, "Gait reference trajectory generation at different walking speeds using LSTM and CNN", Multimedia Tools and Applications, Vol. 82, pp. 33401?33419, 2023

[7] Sravan Kumar Challa, Akhilesh Kumar, Vijay Bhaskar Semwal, and Nidhi Duam, "An optimized-LSTM and RGB-D sensor-based human gait trajectory generator for bipedal robot walking." IEEE Sensors Journal, Vol. 22(24), pp. 24352-24363, 2022

[8] Semwal, Vijay Bhaskar and Gora Chand Nandi. "Toward Developing a Computational Model for Bipedal Push Recovery-A Brief." IEEE Sensors Journal, Vol. 15(4) pp. 2021-2022, 2015

[9] Semwal, Vijay Bhaskar, et al. "An optimized feature selection using biogeography optimization technique for human walking activities recognition", Computing, Vol. 103(12) pp. 2893-2914, 2021

[10] Vijay Bhaskar Semwal, Pratibha TokasPratibha, TokasMaske, and NileshNilay Khare, "Dynamic balancing of bipedal robot on inclined terrain using polynomial trajectories." In Proc. of the 2023 7th International Conference on Graphics and Signal Processing. 2023.

[11] Wenbo Huang, Lei Zhang, Qi Teng, Chaoda Song, and Jun He, "The Convolutional Neural Networks Training With Channel-Selectivity for Human Activity Recognition Based on Sensors," *IEEE Journal of Biomedical and Health Informatics*, Vol. 25(10), pp. 3834 - 3843, 2021

[12] Sakorn Mekruksavanich and Anuchit Jitpattanakul, "Sport-Related Activity Recognition from Wearable Sensors Using Bidirectional GRU Network,"*Intelligent Automation & Soft Computing*, Vol.34(3), pp.1907-1925, 2022

[13] Qi Teng, Kun Wang, Lei Zhang, and Jun He, "The layer-wise training convolutional neural networks using local loss for sensor-based human activity recognition,"*IEEE Sensors J.*, vol. 20, no. 13, pp. 7265–7274, Jul. 2020.

[14] Yin Tang, Qi Teng, Lei Zhang, Fuhong Min, and Jun He, "Layer-wise training convolutional neural networks with smaller filters for human activity recognition using wearable sensors," *IEEE Sensors J.*, vol. 21(1), pp. 581–592, Jan. 2021.

[15] H. Cho and S. Min Yoon, "Divide and Conquer-Based 1D CNN Human Activity Recognition Using Test Data Sharpening," *Sensors*, Vol. 18(4), April 2018.

[16] Jiahui Huang, Shuisheng Lin, Ning Wang, Guanghai Dai, Yuxiang Xie, and Jun Zhou, "TSE-CNN: A Two-Stage End-to-End CNN for Human Activity Recognition," *IEEE Journal of Biomedical and Health Informatics*, Vol. 24(1). 2019.

[17] Myung-Kyu Yi, Wai-Kong Lee and Seong Oun Hwang, A Human Activity Recognition Method Based on Lightweight Feature Extraction Combined with Pruned and Quantized CNN for Wearable Device,"*IEEE Transactions on Consumer Electronics*, 2023.

[18] Weiheng Kong, Lili He, and Hailong Wang, "Exploratory Data Analysis of Human Activity Recognition Based on Smart Phone", *IEEE Access*, Vol.9, pp.73355-73364, 2021.

[19] Nabil Zerrouki, Fouzi Harrou, Ying Sun, and Amrane Houacine, "Vision-Based Human Action Classification Using Adaptive Boosting Algorithm,"*IEEE Sensors Journal*, Vol. 18(12), pp. 5115 - 5121, 2018

[20] Joao A. Francisco and Paulo Sergio Rodrigues, "Computer Vision Based on a Modular Neural Network for Automatic Assessment of Physical Therapy Rehabilitation Activities,"*IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 31, pp. 2174 - 2183, 2022

[21] Heilym Ramirez, Sergio A. Velastin, Ignacio Meza, Ernesto Fabregas, Dimitrios Makris, and Gonzalo Farias, "Fall Detection and Activity Recognition Using Human Skeleton Features,"*IEEE Access*, Vol. 9, pp 33532 - 33542, 2021

[22] Hochul Hwang, Cheongjae Jang, Geonwoo Park, Junghyun Cho, and Ig-Jae Kim ElderSim: A Synthetic Data Generation Platform for Human Action Recognition in Eldercare Applications,"*IEEE Access*, Vol 11, pp. 9279 - 9294, 2021

[23] Muhammad Ehatisham-Ul-Haq, Ali Javedm, Muhammad Awais Azam, Hafiz M. A. Malik, Aun Irtazam, Ik Hyun Lee, and Muhammad Tariq Mahmood, "Robust Human Activity Recognition Using Multimodal Feature-Level Fusion,"*IEEE Access*, Vol. 7, pp. 60736 - 60751, 2019.

[24] B. Barshan and M. C. Yüksek, "Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units,"*The Computer Journal*, Voo. 57(11), pp.1649–1667, November 2014.

[25] Hong-Lam Le, Duc-Nhan Nguyen, Thi-Hau Nguyen,and Ha-Nam Nguyen, "A Novel Feature Set Extraction Based on Accelerometer Sensor Data for Improving the Fall Detection System," *Electronics*, Vol. 11(7), 1030, 2022.

[26] Le Wang, Meng Han, Xiaojuan Li, Ni Zhang, and Haodong Cheng, "Review of Classification Methods on Unbalanced Data Sets,"*IEEE Access*, Vol. (9) pp. 64606 - 64628, 2021

[27] Jesper E. van Engelen and Holger H. Hoos, "A survey on semi-supervised learning," *Machine Learning volume*, Vol. 109, pp. 373–440, 2020.

[28] Hazar Zilelioglu, Ghazaleh Khodabandelou, Abdelghani Chibani, and Yacine Amirat, "Semi-Supervised Generative Adversarial Networks with Temporal Convolutions for Human Activity Recognition,"*IEEE Sensors*, April 2023.

[29] E. Soleimani, G. Khodabandelou, A. Chibani, and Y. Amirat, "Generic semi-supervised adversarial subject translation for sensor-based activity recognition,"*Neurocomputing*, vol. 500, pp. 649–661, Aug. 2022.

[30] Andrea Rosales Sanabria, Franco Zambonelli, and Juan Ye, "Unsupervised Domain Adaptation in Activity Recognition: A GAN-Based Approach," *IEEE Access*, Vol. 9, pp.19421-19438, 2021

[31] Yiqiang Chen, Jindong Wanga, Meiyu Huang, and Han Yu, "Cross-position activity recognition with stratified transfer learning," *Pervasive and Mobile Computing*, Vol. 57, pp. 1–13, 2019.

[32] Jiachen Zhao , Fang Deng , Haibo He , and Jie Chen, "Local Domain Adaptation for Cross-Domain Activity Recognition ," *IEEE Transactions on Human-Machine Systems*, Vol. 51(1), pp. 12-21. February 2021.

[33] Majeed, A., and Hwang, S. O, "Data-centric AI, Pre-Processing, and the Quest for Transformative AI Systems Development,"*Computer*, 56(5), 1-6, 2023

[34] Eliza Strickland, "Andrew Ng, AI Minimalist: The Machine-Learning Pioneer Says Small is the New Big," *IEEE Spectrum*, Vol. 59(4), pp. 22-50, April 2022

[35] Lei Zhang , Hao Wu , Fuhong Min , and Aiguo Song, "Channel-Equalization-HAR: A Light-weight Convolutional Neural Network for Wearable Sensor Based Human Activity Recognition Wenbo Huang ,*IEEE TRANSACTIONS ON MOBILE COMPUTING*, Vol. 22(9), pp.757 - 775, 2022.

[36] Semwal, Vijay Bhaskar, Anjali Gupta, and Praveen Lalwani. "An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition." The Journal of Supercomputing, Vol. 77, pp. 12256-12279, 2021

[37] Dipanwita Thakur, Suparna Biswas, Edmond S. L. Ho, and Samiran Chattopadhyay, "ConvAE-LSTM: Convolutional Autoencoder Long Short-Term Memory Network for Smartphone-Based Human Activity Recognition,"*IEEE Sensors*, Vol. 10, pp. 4137 - 4156, 2022.

[38] Kun Xia, Jianguang Huang, and Hanyu Wang, "LSTM-CNN Architecture for Human Activity Recognition," *IEEE Access*, Vol. 8, pp. 56855 - 56866, March 2020.

[39] Gaud, Neha, Maya Rathore, and Ugrasen Suman, "Hybrid Deep Learning-Based Human Activity Recognition (HAR) Using Wearable Sensors: An Edge Computing Approach," In Proc. of International Conference on Data Analytics & Management. Singapore, Springer Nature Singapore, 2023.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3406727

[40] Manoj Kumar Sain, Joyeeta Singha, Sandeep Saini, and Vijay Bhaskar Semwal, "Human Action Recognition Using ConvBiLSTM-GRU In Indoor Environment," In Proc. of the 2023 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT). 2023.

[41] Sravan Kumar Challa, Akhilesh Kumar, Vijay Bhaskar Semwal, and Nidhi Dua, "An optimized deep learning model for human activity recognition using inertial measurement units." Expert Systems, Vol. 40, pp.1-18, 2022.

[42] Mohammed A. A. Al-qaness , Abdelghani Dahou , Mohamed Abd Elaziz , and A. M. Helmi, "Multi-ResAtt: Multilevel Residual Network With Attention for Human Activity Recognition Using Wearable Sensors*IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, Vol. 19(1), pp. 144 - 152, 2022

[43] Enes Kosar and Billur Barshan, "A new CNN-LSTM architecture for activity recognition employing wearable motion sensor data: Enabling diverse feature extraction,*Engineering Applications of Artificial Intelligence*, Vol 124, 2023,

[44] Sakorn Mekruksavanich, Narit Hnoohom, and Anuchit Jitpattanakul, "A Hybrid Deep Residual Network for Efficient Transitional Activity Recognition Based on Wearable,*Applied Sciences*, Vol.12(10), 2022.

[45] Cunyi Yin, Xiren Miao, Jing Chen, Hao Jiang, Deying Chen, Yixuan Tong, and Shaocong Zhengm, "Human Activity Recognition with Low-Resolution Infrared Array Sensor Using Semi-supervised Cross-domain Neural Networks for Indoor Environment,"*IEEE Internet of Things Journal*, 2023.

[46] Paola Ariza Colpas, Enrico Vicario, Emiro De-La-Hoz-Franco, Marlon Pineres-Melo, Ana Oviedo-Carrascal, and Fulvio Patara, "Unsupervised Human Activity Recognition Using the Clustering Approach: A Review,"*Sensor*,20(9), 2702, 2022.

[47] A. Reiss and D. Stricker, "Introducing a New Benchmarked Dataset for Activity Monitoring", The 16th IEEE International Symposium on Wearable Computers(ISWC), 2012.

[48] Long Wen, Liang Gao, Xinyu Li, and Bing Zeng, "Convolutional Neural Network With Automatic Learning Rate Scheduler for Fault Classification," *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, Article Sequence Number 3509912, 2021.

[49] Florenc Demrozi, Graziano Pravadelli, Azra Bihorac, and Parisa Rashidi, "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey," *IEEE Access*, Vol. 8, pp: 210816 - 210836, November 2020.

[50] Jun-Huai Li, Ling Tian, Huaijun Wang, Yang An, and Kan Wang,Lei Yu, "Segmentation and Recognition of Basic and Transitional Activities for Continuous Physical Human Activity,"*IEEE Access*, Vol. 7, pp. 42565 - 42576, March 2019.

[51] C. Chatzaki, M. Pediaditis, G. Vavoulas and M. Tsiknakis, "Human daily activity and fall recognition using a smartphone's acceleration sensor" in Information and Communication Technologies for Ageing Well and e-Health, Springer, vol. 736, 2017.

[52] Jennifer R. Kwapisz, Gary M. Weiss and Samuel A. Moore, "Activity Recognition using Cell Phone Accelerometers", ACM SigKDD Explorations Newsletter, Vol. 12, No. 2, pp. 74–82, 2011.

[53] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems", Montreal, QC, Canada, pp. 2672–268, 8–13 December 2014.

[54] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, Xi Chen, "Improved Techniques for Training GANs, Part of Advances in Neural Information Processing Systems", pp. 2234–2242, Vol(29),2016

[55] Augustus Odena, "Semi-Supervised Learning with Generative Adversarial Networks," Data Efficient Machine Learning workshop at ICML 2016, 2016.

[56] "TensorFlow Libraries", [Online] Available: `https://www.tensorflow.org/resources/libraries-extensions`

[57] Che, Tong, Li, Yanran, Jacob, Athul Paul, Bengio, Yoshua, and Li, Wenjie, "Mode regularized generative adversarial networks,"In International Conference on Learning Representations(ICLR), 2017.

[58] T. Zebin, P. Scully, N. Peek, A. Casson, and K. Ozanyan, "Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition," *IEEE Access*, vol. 7, pp. 133509-133520, 2019.

MYUNG-KYU YI (Member, IEEE) received a Ph.D. in Computer Science and Engineering from Korea University in 2005. He is currently a Research Professor with Gachon University, and is a member of the personal health information standardization task force of the TTA U-Health project group. His research interests include healthcare, security, machine learning, deep learning, and human activity recognition.

SEONG OUN HWANG (Senior Member, IEEE) received the B.S. degree in mathematics from Seoul National University, in 1993, the M.S. degree in information and communications engineering from the Pohang University of Science and Technology, in 1998, and the Ph.D. degree in computer science from the Korea Advanced Institute of Science and Technology, South Korea. He worked as a Software Engineer with LG-CNS Systems, Inc., from 1994 to 1996. He also worked as a Senior Researcher with the Electronics and Telecommunications Research Institute (ETRI), from 1998 to 2007. He worked as a Professor with the Department of Software and Communications Engineering, Hongik University, from 2008 to 2019. He is currently a Professor with the Department of Computer Engineering, Gachon University. His research interests include cryptography, cybersecurity, and artificial intelligence. He is an Editor of ETRI Journal.