Optimizing Sensor and Data Selection on Lower Limbs via Deep Learning for Real-Time Human Activity Recognition

Zihang You, Neethan Ratnakumar, *Member, IEEE*, Bo Shen, *Member, IEEE*, Hao Su, *Senior Member, IEEE*, and Xianlian Zhou, *Member, IEEE*

***Abstract*—Real-time human activity recognition (HAR) is crucial for adaptive control in lower limb exoskeletons, yet achieving an optimal balance between accuracy, latency, and sensor complexity remains challenging. This study's key idea is to systematically evaluate sensor selection and data modalities for real-time HAR using deep neural networks (DNNs) with ultra-short 50 ms sliding windows. Leveraging a dataset from 21 subjects performing six locomotion activities and employing three deep learning models [multilayer perceptron (MLP), LSTM, CNN-LSTM], we assess multiple sensor combinations, including joint angle data, derived angular velocities, and inertial measurements, and quantify their trade-offs across accuracy, latency, and model complexity. Our analysis reveals that bilateral joint angles from hip, knee, and ankle achieve 98.98% accuracy, significantly outperforming unilateral sensor setups. Adding a thigh-mounted inertial measurement unit further elevates accuracy to 99.23%, highlighting the advantages of multimodal sensor fusion. In addition, incorporating derived joint angular velocities enhances accuracy, with up to 15% increase when using single-joint bilateral inverse kinematics data. Even a minimal configuration, such as bilateral hip joints with derived angular velocities, achieves over 94% accuracy, offering practical solutions for low-power wearable systems. These findings establish actionable design principles for HAR-driven control in assistive robotics and mobile health applications.**

***Index Terms*—Deep learning, human activity recognition (HAR), locomotion mode, neural network, sensor selection.**

I. Introduction

HUMAN lower limb exoskeletons have made significant strides in rehabilitation and mobility augmentation. However, effectively integrating these devices with users' natural motion remains challenging. Conventional control methods, which rely on preprogrammed motion trajectories or direct actuator control, often struggle to adapt when transitioning between locomotion modes, such as walking, climbing stairs, or standing. To address this limitation, recent research has focused on developing precise control strategies for specific movements [1], [2], [3]. Human activity recognition (HAR) enables real-time intent detection, allowing exoskeletons to switch between these movement-specific controllers automatically [4], [5]. This eliminates manual mode selection, reduces cognitive load, and enhances usability in dynamic environments [6]. High accuracy HAR is typically associated with appropriate sensor type and placement selection. An optimized sensor configuration can reduce data redundancy, lower system complexity and latency, and establish a strong foundation for adaptive control for both high-level in lower-limb exoskeleton systems [7], [8].

In this study, we present a systematic evaluation of sensor and data selection for real-time HAR. Utilizing a comprehensive human activity dataset [9], we systematically assess sensor modalities commonly used in HAR, their corresponding feature channels, and placement options. Our objective is to balance sensor quantity and types while ensuring high HAR accuracy. This design improves sensor configuration and inference efficiency and helps mitigate challenges related to sensor synchronization, computational complexity, and potential data loss in real-world exoskeleton applications. Distinct from prior work [4], [10], [11], we explore the use of ultra-short (50 ms) time windows as inputs to deep learning models, enabling real-time HAR without sacrificing accuracy. Furthermore, we investigate the impact of derived feature channels (e.g., joint velocity) on model performance, showing their effectiveness in compensating for reduced physical sensing. The use of such synthetic data may prove advantageous, particularly when reducing the number of physical sensors.

II. Related Work

HAR is a well-established challenge in machine learning, with numerous approaches developed over the years. These methods can be broadly categorized into sensor-based and nonsensor-based techniques. Nonsensor approaches include those relying on computer vision [12] and Wi–Fi-based techniques [13],

Received 24 July 2025; revised 21 October 2025 and 2 December 2025; accepted 1 January 2026. The work of Hao Su was supported in part by NSF under Grant CPS 2344956 and in part by the National Institutes of Health (NIH) under Grant 1R01EB035404-01. This work was supported in part by the National Science Foundation (NSF) under Grant GCR-2524088, and in part by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) under Grant 90REGE0025. This article was recommended by Associate Editor C. Lv. *(Corresponding author: Xianlian Zhou.)*

Zihang You, Neethan Ratnakumar, and Xianlian Zhou are with the Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: zy4@njit.edu; nr473@njit.edu; alexzhou@njit.edu).

Bo Shen is with the Department of Mechanical and Industrial Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: bo.shen@njit.edu).

Hao Su is with the Department of Biomedical Engineering, Tandon School of Engineering, New York University, Brooklyn, NY 11201 USA (e-mail: hao.su@nyu.edu).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/THMS.2026.3651525.

Digital Object Identifier 10.1109/THMS.2026.3651525

[14], whereas sensor-based methods utilize wearable sensors to capture motion data. The sensor-based approaches have been widely adopted for distinguishing daily human activities (such as sitting, standing, walking, and stair climbing) [15], with applications in healthcare, sports, and assistive technologies. Commercial wearables, such as smartphones [16], utilize inertial measurement units (IMUs) for basic activity recognition. Foundational datasets, such as the University of California, Irvine (UCI) HAR study [17] demonstrated that minimal sensor configurations such as these can be sufficient to classify stationary activities (e.g., sitting and lying) and some basic dynamic movements (e.g., walking) with reasonable accuracy.

However, distinguishing more nuanced movement variations, particularly within locomotion modes, requires careful sensor selection. Optimizing sensor selection in HAR is crucial for improving wearability and reducing power consumption without compromising accuracy. Zappi et al. [18] analyzed the tradeoff between sensor power consumption and HAR accuracy and Yu et al. [19] reviewed commonly used sensor types and their placement strategies. However, their work primarily provided guidance on sensor choice rather than assessing how sensor quantity and placement together impact classification accuracy. Prior studies provide guidance on sensor choice but do not jointly assess how sensor quantity and placement affect HAR accuracy, and most pipelines rely on longer windows tied to gait-cycle segmentation, which is incompatible with real-time exoskeleton control. This work addresses this gap by systematically evaluating sensor types, placements, and derived channels for 50 ms real-time HAR with DNNs. Camargo et al. [10] applied traditional machine learning and multilayer perceptron (MLP) models on extracted features to classify ambulation modes and estimate terrain parameters (e.g., stair height, ramp incline), assessing the impact of redundancy on classification accuracy and reinforcing the need for optimized sensor selection in HAR. Beyond sensor selection, identifying the most informative feature channels within each sensor is equally important, particularly in time-series HAR [20]. An IMU typically includes a tri-axial accelerometer and a gyroscope, yielding six distinct feature channels of acceleration and angular velocity. In contrast, electronic goniometers are limited to sagittal and frontal plane measurements but offer direct joint angle estimation. Despite providing fewer feature channels, joint angles have been shown to be highly informative for HAR. The work of Lee et al. [21] discusses the impact of using goniometer data from the unilateral and bilateral knee and ankle joints on the performance of HAR models. Similarly, Kang et al. [22] used exoskeleton joint encoders derived from inverse kinematics (IK) data and found that excluding hip joint angles significantly reduced HAR performance. In the context of human lower limb exoskeletons [23], encoders embedded within actuators are a more commonly utilized means of angle data acquisition. Advancements in sensor selection frameworks, such as automated sensor-feature subset optimization [24], have demonstrated the ability to dynamically reduce sensor redundancy while addressing complex recognition challenges, leading to improved accuracy in classification and estimation tasks.

Real-time HAR systems face critical design trade-offs between temporal resolution and model complexity. While traditional machine learning methods [support vector machine (SVM), random forests, linear discriminant analysis (LDA)] excel at feature selection and data compression for efficient training [25], [26], their reliance on gait cycle segmentation (typically 150–200 ms) [4] introduces latency incompatible with exoskeleton control requirements. This limitation motivates the adoption of deep learning architectures capable of processing shorter time windows (< 150 ms) without compromising temporal dependencies. Long Short-Term Memory (LSTM) networks address this challenge through memory cells and gating mechanisms, and they are particularly effective in modeling sequential data, making them well-suited for HAR tasks where the temporal dynamics of sensor data are crucial [27], while CNN-LSTM hybrids combine convolutional layers for spatial feature extraction with LSTM temporal modeling [28]. These architectures achieve state-of-the-art performance on benchmark datasets (95.85% on WISDM [29], 92.1% on LSTM implementations [30]), outperforming both traditional methods and basic MLP that serve as nontemporal baselines [17]. The choice between these approaches involves fundamental trade-offs: shallow models (MLPs) provide computational efficiency for well-defined feature spaces, whereas LSTM variants capture complex activity patterns through inherent temporal reasoning, a critical capability for real-time exoskeleton control. Beyond these, emerging deep learning models continue to push HAR performance. For example, Ronald et al. introduced the iSPLInception model based on a ResNet architecture, achieving 95% accuracy on the UCI-HAR dataset [31], while transformer-based models leverage self-attention mechanisms to capture long-range temporal dependencies [32]. These advances highlight accuracy gains, but also reinforce the persistent trade-off between model complexity and real-time feasibility.

Although HAR is a well explored area, significant challenges persist in determining optimal sensor configurations for various applications. Numerous studies have focused on optimizing IMU placement under various conditions. The work by Xia et al. [33] established an optimization scheme for virtual sensor data to determine the optimal placement of IMUs for different activities; Zheng et al. [34] proposed a method for human posture reconstruction achieved by optimizing the selection of IMUs and using bidirectional recurrent neural networks. Caramia et al. [35] conducted an analysis of optimized IMU configurations in the assessment of Parkinson's Disease. The study by Poulose et al. [36] investigated the performance of IMU sensors placed on the wrist, waist, and ankle across various physical activities. However, identifying a universally optimal HAR setup remains a challenge due to varying experiment designs, sensor types, and configurations of machine learning or deep learning models [37], [38]. Real-time HAR further adds another layer of complexity, often requiring a reduction in the amount of data processed by deep learning models along the time dimension while maintaining accuracy.

III. Methods

A. Data Selection

Our sensor and data selection strategy involved systematically evaluating various combinations of sensor types, anatomical placements, and derived features available in a public biomechanics dataset. These configurations were assessed using multiple deep learning models to identify minimalist sensor sets that enable high-accuracy, low-latency HAR suitable for wearable assistive systems. For this purpose, the study utilized the Georgia Tech lower limb biomechanics dataset [9] comprising 21 subjects (out of 22, excluding subject AB07 due to data-label desynchronization). Each subject was equipped with 3 goniometers (hip, knee and ankle), 4 six-axis IMUs on the right lower limb (thigh, shank and foot) and trunk, and 32 motion capture

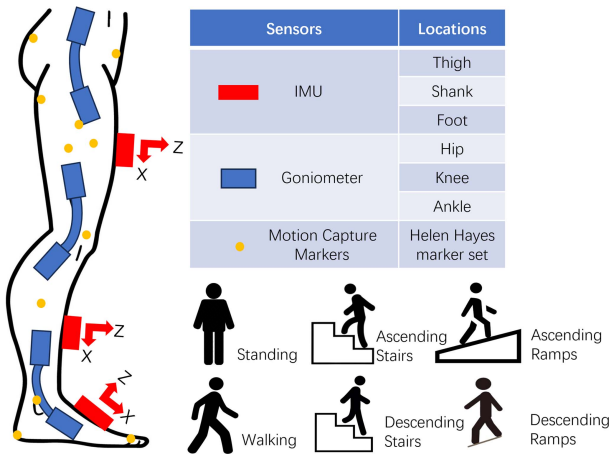


Fig. 1. Placement of utilized IMUs, goniometers, and motion capture markers. More details on sensor and marker placement can be found in [9].

markers based on the Helen Hayes Hospital marker set [39]. The data utilized in our study comprises of the bilateral flexion angles of the hips, knees, and ankles, computed through IK on the marker data. While marker-based IK data are not directly applicable to real-time exoskeleton control, they were used to simulate joint angle data provided by encoders, considering the similarity between IK joints and encoders in providing joint angle data. This also facilitates the comparison of bilateral (IK) and unilateral joint angles (goniometers) in HAR performance. Although the dataset includes electromyographic (EMG) signals, which are valuable for muscle activation analysis, we excluded EMG data from our analysis due to its inherent noise and variability [40]. EMG signals typically require precise electrode placement and skin contact, which introduce practical challenges in daily exoskeleton use. The placement of utilized IMUs, goniometers, and motion capture markers is shown in Fig. 1. All the sensors and motion capture markers are directly mounted on the human body. Given our focus on lower limb-mounted sensors for exoskeleton integration, IMU data from the trunk was excluded to better reflect the scenario of sensor integration in lower limb exoskeletons. For IMU data, we retained all six axes of accelerometer and gyroscope measurements as input to the models. In the original dataset, goniometers were used to measure the joint angles in the frontal and sagittal planes of the hip and ankle, and only in the sagittal plane of the knee joint. We restricted our analysis to sagittal plane goniometer data for all joints (hip, knee, ankle), as sagittal kinematics (e.g., hip/knee flexion/extension) are biomechanically dominant in the studied locomotion activities. Similarly, from the IK joint angle data, only the sagittal plane angles of the hip, knee, and ankle joints were utilized. In the dataset, the IMU data was sampled at 200 Hz and postprocessed using a 6th-order Butterworth low-pass filter with a cutoff frequency of 100 Hz. The goniometer data was sampled at 1000 Hz and processed using a 4th-order Butterworth low-pass filter with a cutoff frequency of 20 Hz. The IK joint angle data available in the dataset was also given at 200 Hz and postprocessed using a 4th-order zero-lag Butterworth low-pass filter with a cutoff frequency of 6 Hz. For consistence, we downsampled the goniometer time series at 200 Hz, aligning with the sampling frequency of both the IMU and IK data.

The dataset provides more than 20 types of activity labels, and we selected 6 common and representative action labels: standing (ST), walking on level ground (WL) at self-selected speeds (0.88–1.45 m/s), ascending/descending on ramps (RA/RD) across six inclines (5.2°–18°), and ascending/descending on stairs (SA/SD) with four step heights (4–7 in.). Each subject completed 30 WL trials, 40 SA/SD trials, and 60 RA/RD trials, yielding approximately 15 min of steady-state kinematic data per participant.

B. Data Post-Processing

A sliding window segmentation method was employed to transform time series data into tensor formats suitable for training deep learning models. The window size and the overlap rate are the two key parameters that affect the input of the HAR models. For algorithms primarily focused on successfully identifying activities, the maximum sliding window duration can be as long as 10 s [41]. However, for algorithms requiring quick response, such as those used in exoskeleton control, the sliding window duration can be much shorter (e.g. 150 ms [4]), provided that accuracy is satisfactory. In our implementation, the length of the sliding window is set at 50 ms with an overlap of 50%. We fixed the overlap at 50% across all configurations to ensure temporal continuity between consecutive windows and minimize redundancy in real-time processing. A 50% sliding window overlap was also commonly used in other HAR studies [42], [43], [44].

Sikder et al. [45] indicated that increasing the number of feature channels in NN models can enhance the model's ability to learn complex and diverse features, which in turn can improve the accuracy and robustness of HAR models. When using only goniometers, joint angles are the sole input features available for the HAR model. While joint angular velocities can be directly measured using IMUs [46], [47], such approaches necessitate additional hardware, increasing system complexity and cost. To circumvent this, we approximate the joint angular velocities by differentiating joint angle data using the central difference method, eliminating the need for extra sensors. This approach was applied to angular data from both the goniometer and IK joint angles to derive the sagittal plane angular velocities of various joints, offering a practical solution in scenarios where direct angular velocity measurements are unavailable or impractical.

C. Neural Networks

We implemented three DNN models introduced in Section I: MLP, LSTM, and CNN-LSTM. Depicted in Fig. 2(a), our LSTM model comprises one LSTM layer with 100 neurons, a 50% dropout layer, and two dense, fully connected layers. The MLP architecture [see Fig. 2(b)] comprised of an input layer, a hidden layer (with 128 neurons), and an output layer. The CNN part of the CNN-LSTM integrates two convolutional layers with a kernel size of 3, an additional 50% dropout layer, and a pooling layer. Fig. 2(c) illustrates the CNN component's architecture. An additional flattened layer is located at the end to compress the dimensions of the CNN output tensor to match the dimensions of the input required by the LSTM. The subsequent LSTM structure attached to the CNN mirrors that of the standalone LSTM.

For training and testing of the deep learning models, each subject's data after the sliding window method was divided into training and test sets with a ratio of 7:3 and the subject-specific models were trained for each subject individually. In our experimental setup, we employed the Adam optimizer with cross-entropy as the loss function across all three model architectures. For hyperparameter setting, we maintained a consistent

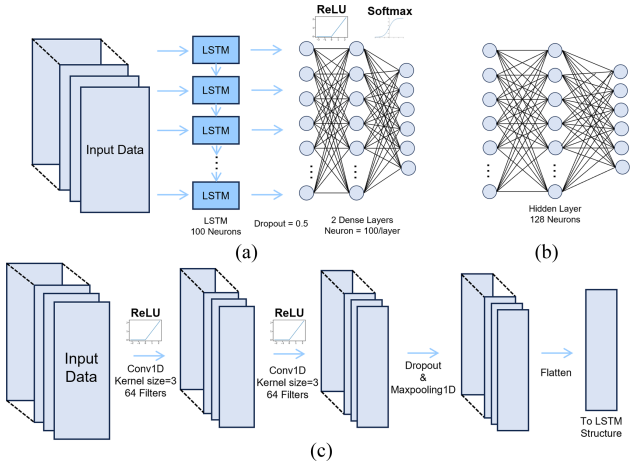


Fig. 2. (a) Structure of the LSTM network. (b) The structure of the MLP network. (c) The structure of the CNN part of the CNN-LSTM network.

batch size of 64. The number of training epochs were varied by model type to ensure proper convergence for all subjects: 15 epochs for LSTM models, 20 epochs for MLP models, and 25 epochs for CNN-LSTM models. To enhance statistical robustness, we conducted 20 independent testing rounds for each sensor configuration on each subject. In each round, the training and testing data of the subject were shuffled at the beginning for the training of a new model. During the model training iterations and followed testing, the original train-test split was preserved. This approach guarantees that the model is exposed only to training data during fitting, while testing data remains unseen until final evaluation. Each round of training generated a new model and we obtained the mean accuracy from 20 rounds of testing for each subject. Subsequently, we used these means to compute the mean and standard deviation accuracy for all subjects. This repeated random subsampling cross-validation approach reduces the randomness associated with any single data split and enables the calculation of both the mean and variance of model performance, resulting in a more stable and statistically meaningful evaluation. During our evaluation of these three deep learning models, we did not perform systemic hyperparameter tuning for each sensor configuration. This allows a more direct comparison of how different sensor configurations impact HAR results without introducing uncertainty of tuning.

During training, the cross-entropy loss for all three models dropped sharply within the first few epochs, then decreased gradually before leveling off. By the end of training, each model converged to a low loss value, indicating effective learning and confirming that the model architectures and optimizers were appropriately configured.

D. Inertial Measurement Units

IV. Results

We evaluated the performance of all possible combinations of the three IMUs, including individual IMUs, any pair of IMUs, and all three IMUs together. All six axes of IMU accelerometer and gyroscope data were used for training. The mean classification accuracy and standard deviation, averaged for all 21 subjects, are present in Fig. 3. The results reveal that using two or more IMUs on one side of the body achieves HAR accuracy consistently above 95%. While reducing the number

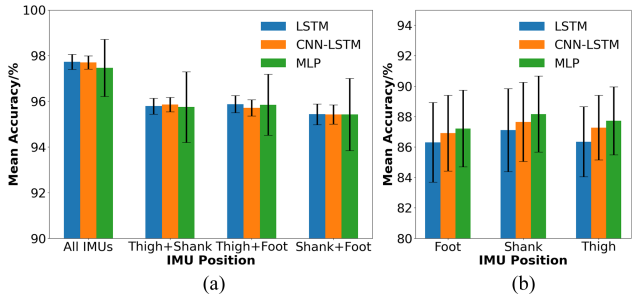


Fig. 3. (a) HAR accuracy with all three IMUs and any pair of IMUs. (b) HAR accuracy with individual IMUs.

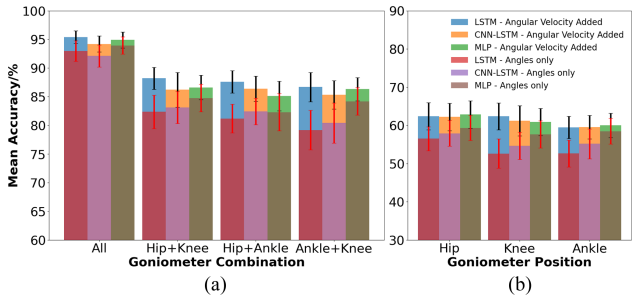


Fig. 4. (a) HAR accuracy with all three goniometers and any pair of goniometers. (b) HAR accuracy with individual goniometers.

of IMUs in HAR models leads to a decline in accuracy, it is noteworthy that any individual IMU still maintains a relatively good average accuracy, exceeding 86% Statistical analysis using repeated-measures ANOVA (rm-ANOVA) confirmed that increasing the number of IMUs, significantly enhances HAR accuracy (all $p < 0.05$). Posthoc paired *t*-test with Bonferroni correction ($\alpha = 0.0167$, all $p < 0.001$) also suggests that the number of features in the dataset plays a crucial role. However, a post-hoc paired *t*-test revealed that the placements of the IMU does not significantly impact on HAR accuracy ($p > 0.083$). As shown in Fig. 3(a), LSTM and CNN-LSTM models exhibit smaller standard deviations compared to MLP (e.g., 0.33 for LSTM, 0.28 for CNN-LSTM, vs. 1.25 for MLP, with all IMUs), indicating superior stability across sensor configurations.

A. Goniometer

Similar combination tests were conducted on the HAR performance using goniometers. The mean classification accuracy and standard deviation, averaged for all 21 subjects, are present in Fig. 4. Overall, the results indicate that increasing the number of goniometers used improves the accuracy of the HAR model. Using only the time series of the angles provided by the goniometers, the HAR accuracy with two goniometers on one side of the body can reach between 80% to 85%, and up to 95% when three goniometers were applied. In contrast, using only a single goniometer resulted in a significant lower accuracy, ranging from 50% to 65%. rm-ANOVA analysis confirmed that increasing the number of goniometers significantly improves the HAR accuracy (all $p < 0.05$), with the post-hoc paired *t*-test, with Bonferroni correction ($\alpha = 0.0203$, all $p < 0.001$), suggesting the number of input data channels is a key factor. However, for tests involving one or two goniometers, a post-hoc paired *t*-test revealed that the specific choice of goniometers does not significantly impact HAR accuracy ($p > 0.084$). In addition,

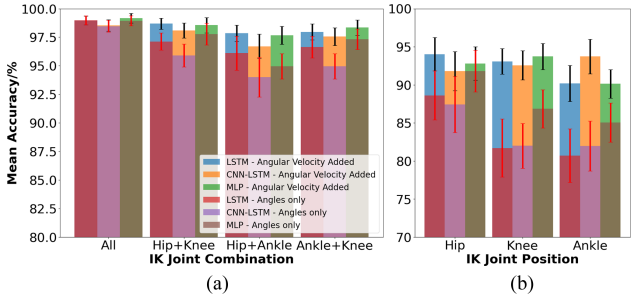


Fig. 5. (a) Bilateral IK HAR accuracy with all three joints and any pair of joints. (b) Bilateral IK HAR accuracy with individual joints.

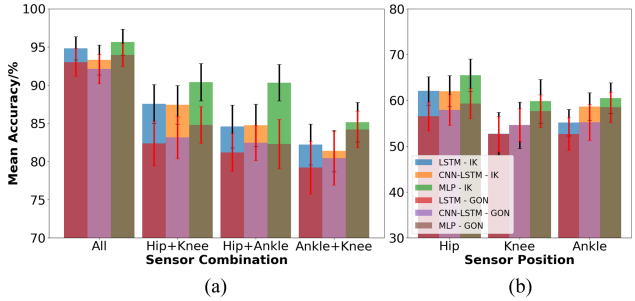


Fig. 6. (a) HAR performance comparison between goniometers and IK (joint angle only) with all three joints and any pair of joints. (b) HAR performance comparison between goniometers and IK (joint angle only) with individual joints.

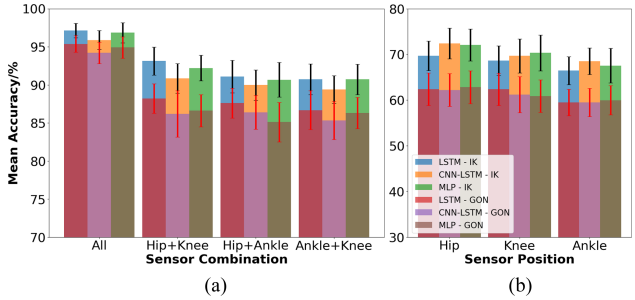


Fig. 7. (a) HAR performance comparison between goniometers and IK (with angular velocity) with all three joints and combinations of any two joints. (b) HAR performance comparison between goniometers and IK (with angular velocity) with a single joint.

the inclusion of angular velocity, calculated from goniometer joint angles with the central difference method, led to a significant increase of HAR accuracy across all three DNN models (paired t-test, all $p < 0.05$). In the scenario where all three joints are utilized, incorporating joint angular velocities enhances the performance of the three models by approximately 1% –3% Notably, the LSTM model showed the greatest improvement, nearly 10%, when angular velocities were added to the "ankle+knee" combination, followed by the CNN-LSTM model, with the MLP model showing the least improvement.

B. Inverse Kinematics

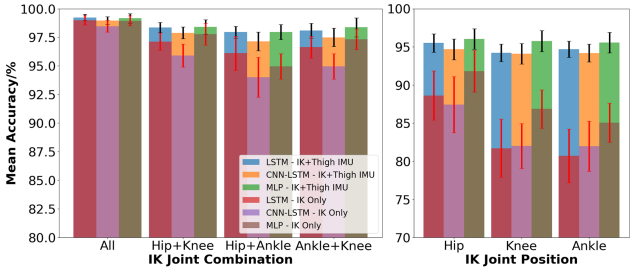
In the study by Yu et al. [4], HAR using bilateral thigh IMU motion data was applied to a hip exoskeleton, achieving excellent results on four activities (standing, level walking, and ascending/descending stairs). Given that the dataset we use did not include IMUs mounted on both lower limbs, we utilized bilateral hip joint IK data and generated hip angular velocities with three DNN models to test the same four activities. The IK data of the GT dataset provides bilateral lower limb joint angles (six in total) calculated from the motion capture data. Similar to the goniometers, we also used the central difference method to calculate the angular velocities of all the IK joint angles during the activities. The mean classification accuracy and standard deviation averaged for all 21 subjects are shown in Fig. 5. The results demonstrate that using bilateral joint data significantly outperform IMUs and goniometers, regardless of the joint combination. When using three joints from both lower limbs, the average accuracy of all three models exceeded 98%, with the highest accuracy of 98.98% using the LSTM model. Due to its high accuracy, adding angular velocity to the joint angle dataset did not further improve the model's performance. For combinations of any two joints, the accuracy of the "hip+ankle" combination is slightly lower than the other two ("hip+knee" and "knee+ankle"). Nonetheless, even the least accurate model in this case, CNN-LSTM, achieved an accuracy above 93% using only "hip+ankle" joint angles. The highest average accuracy was achieved with the "hip+knee" combination, where the accuracy of MLP reached 97.77% with joint angle alone. Incorporating angular velocity in the two joint cases improved the average accuracy across all three models by approximately 2% with the LSTM model seeing the most improvement and reaching 98.56% accuracy with the "hip+knee" configuration. Incorporating angular velocity to individual joints resulted in substantial enhancement in HAR performance. When utilizing only a single joint, the hip exhibits the best accuracy of 91.79% using the MLP, followed by the knee and ankle joint with accuracy between 80% and 87% The inclusion of angular velocity elevates the HAR performance of all individual joints above 90% Notably, adding angular velocity to the knee joint resulted in a performance elevation of nearly 15% for LSTM model. When combined with derived angular velocities, the hip achieves the best accuracy of single joint accuracy at 94.02% using the LSTM.

C. IK vs Goniometers

Given that both IK and goniometers measure the angular movement of joints, we conducted a comparative analysis of the HAR performance using these two sensing modalities. Since the goniometer data was exclusively from the right lower limb, we selected corresponding data from the right lower limb in the IK dataset for a fair comparison. Fig. 6 presents the results of the comparison utilizing only joint angles, while Fig. 7 incorporates the joint angular velocity into the comparison. The comparative results demonstrate that, when relying solely on joint angles, IK data generally outperformed those using goniometers, with IK performance for "hip+knee" and "hip+ankle" combinations leading goniometers by 5% when using the MLP model. Upon the integration of angular velocity, the HAR performance of IK consistently outperforms the goniometer. In scenarios involving all joints, IK's average HAR accuracy surpassed that of the goniometer by 2% When comparing any two joints, IK's advantage increased to 5% and for single-joint tests, the advantage nearly reached 10%

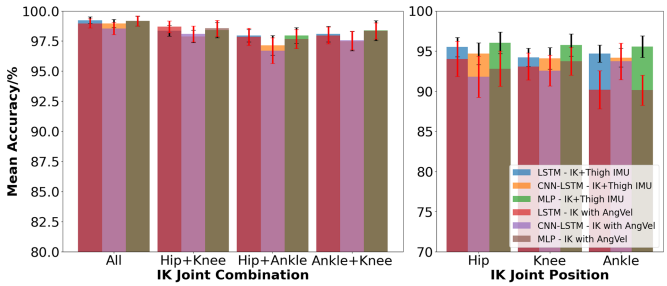
D. IK + Thigh IMU

To mimic a potential sensor configuration for a human hip exoskeleton, we evaluated the combination of bilateral hip sagittal



(a) (b)

Fig. 8. (a) HAR performance comparison between bilateral IK (joint angle only) and IK+Thigh IMU with all three joints and combinations of any two joints. (b) HAR performance comparison between IK (joint angle only) and IK+Thigh IMU with a single joint.



(a) (b)

Fig. 9. (a) HAR performance comparison between IK (with derived angular velocity) and IK+Thigh IMU with all three joints and combinations of any two joints. (b) HAR performance comparison between IK (with derived angular velocity) and IK+Thigh IMU with a single joint.

angle (IK) data with an additional thigh-mounted IMU. Incorporating this additional thigh IMU enhanced the performance of the three models by 6.89% 7.26% and 4.24% respectively, when compared to models with only bilateral hip joint angles. We further conducted analysis by adding the IMU to all joint IK combinations. In Fig. 8, the HAR performance is illustrated both prior to and subsequent to the incorporation of the thigh IMU, utilizing various joint combinations from the bilateral IK dataset. For any joint combination within the IK dataset, inclusion of the thigh IMU enhances the HAR performance across all DNN models. When all joints are included, incorporating the thigh IMU enables the models to maintain robust performance, achieving an average accuracy of approximately 99% across all three models, with the highest accuracy reaching 99.23% For any dual-joint combination, the introduction of the thigh IMU enhances the average accuracy across the three configurations, with the mean accuracy of all models approximating 98% Notably, the "hip+ankle" combination provides the highest rise, roughly 3% achieving an accuracy of 97.96% (MLP). The highest accuracy is brought by the "hip+knee" combination, which reaches an accuracy of 98.41% (MLP). The distinctions among the various combinations remain small. In contrast, integrating the thigh IMU elevated the HAR accuracy of individual joints by approximately 7% –15% converging the HAR performance of all individual joints to around 95% with the best result 96.03% achieved by hip (MLP), followed by 95.74% with knee (MLP). In addition, incorporating the thigh IMU decreased the standard deviation of the mean HAR accuracy, rendering the model's performance more consistent.

In Fig. 9, we compare the performance of bilateral IK and Thigh IMU with bilateral IK and derived joint angular velocity. Overall, the inclusion of the thigh IMU continues to bolster performance in individual joint evaluations, although the margin of enhancement lowers to a range of 1% –5% When employing any combination of two joints or all joints, the contribution of the thigh IMU offers negligible impact on performance. However, compared to using joint angles and angular velocities alone, integrating the thigh IMU reduces the standard deviation of mean accuracy, rendering more consistent model performance.

Using the best performing "bilateral IK + thigh IMU" data combination, we deployed the corresponding deep learning models onto a standalone Nvidia Jetson Xavier (an edge device suited for exoskeleton control) to evaluate their activity prediction speed. To emulate real-world HAR, the test data was fed to the models in single batches, performing real-time recognition for each sliding window. Among the models, the MLP model achieved the fastest average inference time, ranging from 0.43 to 0.71 ms for different joint configurations. The CNN-LSTM model demonstrated a slightly lower inference time (2.84–2.93 ms) than the LSTM model (3.18–3.28 ms), which could potentially be attributed to the GPU's parallel processing efficiency with the CNN-LSTM's architecture. Notably, both the LSTM and CNN-LSTM models maintained inference times under 3.3 ms, suggesting a potential HAR rate exceeding 300 Hz.

E. Confusion Matrices

To further ascertain the capability of sensors to distinguish between these six activities and understand the nature of misclassifications, we generated confusion matrices during testing based on better-performing sensors and sensor combinations of IK and IK+thigh IMU. We first selected the best-performing (i.e., highest average accuracy across subjects) model type of the three DNNs for each sensor configuration based on previous bar chart results. Subsequently, for that particular model type, we chose the model with the highest accuracy (among 20 rounds of individual training and testing) for each subject and averaged the corresponding confusion matrices across all the subjects to obtain the averaged confusion matrix results.

Fig. S1 in the supplement material shows the best classification results from the top performing models using bilateral IK joint angle data, bilateral IK joint angle and angular velocity data, and bilateral IK joint angle data combined with the thigh IMU data, respectively. The confusion matrix for IK indicates that its primary misclassification errors occur between RA and stair walking, as well as level ground walking and ramp walking. As the number of joints decreases, the misclassification rate of SA being mistaken for RA increases from 0.21% (2 joints) to 6.57% (1 joint), and ramp descending misclassified as level ground walking increases from 1.51% to 10.05% With the addition of generated IK joint angular velocities, the error rates for these two misclassifications decrease to 4.09% and 7.79% respectively, when using only a single joint. In optimal conditions, the performance of bilateral IK with angular velocities is very close to that of bilateral IK+thigh IMU. When using three joints, no misclassification rate exceeds 0.05% ; with one joint, the misclassification rate of level ground walking being mistaken for ramp descent is 3.1% and SA being mistaken for ramp ascent is 3.15% The combination of bilateral IK and thigh IMUs performed the best, with all misclassification rates below 1% when using three and two joints, except between ramp descending and level ground walking, which is 1.02% This is primarily due to the absence of misclassification between ramp and stair walking in the same direction. Although this

TABLE I
COMPARISON WITH OTHER DEEP LEARNING TECHNIQUES FOR ACTIVITIES CLASSIFICATION

Ref. Year	Model	Sensors	Accuracy/F1-score	Activities
2020 [50]	CNN	1 IMU	86.7–96.7%	5
2021 [51]	LSTM	6 IMU	>95%	5
2017 [48]	CNN-SVM	1 IMU	94.49%	6
2018 [49]	LSTM	2 IMU	96%	6
2020 [52]	WaveNet	2 IMU	97.88% (F1-score)	7
2021 [53]	CNN	Several IMUs	98% (F1-score)	9
Our Work 1	LSTM	2 hip IK joint angles & derived angular velocity	95.1%	6
Our Work 2	MLP	2 hip IK joint angles & 1 IMU	99.13%	6

error reoccurred when using a single joint, its severity was significantly reduced compared to using bilateral IK alone, with maximum misclassification rate around 3% Among all the joint combinations shown in Fig. S1, the combination of the hip and knee joints exhibited the best performance when using any pair of joints; the hip joint was the best performer when using a single joint. In terms of DNN selection, MLP generally outperformed LSTM and CNN-LSTM in most scenarios. However, LSTM showed slightly better performance than MLP when all three joints were used. Across all types of data, the average accuracy gradually decreased as the number of joints used was reduced, with a more significant drop in accuracy when reducing from two joints to one joint.

We compared our sensor configurations and accuracy results with other similar works, and the results are shown in Table I. When utilizing bilateral hip joint angles (IK) and the derived angular velocities, our results are slightly inferior to [48] and [49], both of which use two IMU sensors. When we use the bilateral angles of the two hip joints and a thigh IMU, our results are better than all other referenced works.

F. Sliding Window Size

To evaluate the impact of different sliding window sizes (durations) on HAR model performance, we conducted tests using near optimal sensor combinations, which includes bilateral IK and the thigh IMU. For consistency across all window durations, we maintained the fixed 50% sliding window overlapping. During these tests, we increased the sliding window sizes to 200 and 300 ms, while reducing the sampling frequency to 50 Hz. Longer windows capture extended temporal context, potentially improving accuracy for activities, but at the expense of higher latency. This adjustment ensures direct comparable window dimensions without the need to modify the model architecture or introduce computational bias from variable input dimensions. Although down sampling may theoretically omit high-frequency signals, a sampling rate at 20–50 Hz is sufficient to capture the kinematic patterns of human daily activities [54], [55]. Fig. 10 illustrates the HAR accuracy of bilateral IK and thigh IMU across various numbers of joints, under sliding window durations of 200–300 ms. When the sliding window length was increased from 50 to 200 ms, the performance of models using three joints remained nearly unchanged, those using two joint combinations showed a performance increase of no more than 0.5% and those using a single joint saw an increase in performance ranging from 0.7% to 2% The 300 ms sliding window setting yielded slightly better performance improvements than the 200 ms setting. Compared to the 50 ms setting, the performance of models using three joints improved by 0.1% those using two joint combinations showed a performance increase of no more than 1% and those using a single joint experienced a performance increase ranging from 1.5% to 3%.

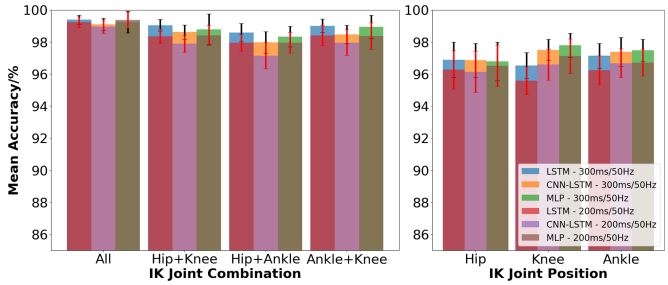


Fig. 10. (a) Bilateral IK+thigh IMU HAR performance comparison under 200 and 300 ms with all three joints and any pair of joints. (b) Bilateral IK+thigh IMU HAR performance comparison under 200 and 300 ms with individual joints.

V. DISCUSSION

This study examines the HAR performance of three distinct sensor types when used individually or in combinations at different placements, aiming to identify the optimal combination of sensors for HAR applications in assistive devices such as lower limb human exoskeletons. We chose an open source dataset that offers a multitude of sensor feature channels, varieties of human activity, and diverse environment conditions to ensure that our models possess a degree of generalizability for future developments on HAR for activities of daily living.

In this study, we adopted a repeated, stratified group shuffle–split scheme per subject rather than a conventional k-fold cross-validation protocol. This procedure can be viewed as a Monte Carlo cross-validation strategy. By repeating the stratified group shuffle–split 20 times and averaging across repetitions, we reduce the sensitivity to any single data split and can estimate both the mean and the variance of model performance. We note that this approach is not inherently more sample-efficient than k-fold cross-validation, and its reliability depends on the number of repetitions and the representativeness of each random split. In our setting, approximately balanced activity classes and large per-subject sample sizes make stratified random partitions statistically close to the underlying subject-level distribution, so repeated Monte Carlo cross-validation provides a practical trade-off between computational cost and evaluation stability.

We concentrated on HAR using a very short sliding time window of 50 ms, targeting fast or real-time applications in exoskeleton assistance. While the impact of sliding window size on HAR for locomotion-based activities has been studied, with medium range window sizes (typically 100–500 ms) often resulting in better accuracy [10], [22], the optimal sliding window size can vary depending on the model and applications. For example, Kang et al. [22] found that the optimal sliding window size is 350 ms for SVM and 750 ms for LDA. In contrast, we opted for a shorter sliding window (50 ms), diverging from the longer windows used in other studies (100 ms [11], 150 ms [4], 50–500 ms [10]). While this choice may result in a slight decrease in accuracy, it significantly shortens response time, promising for real-time exoskeleton operation, and, with the same deep learning models, still achieves comparable performance (within 1%) to longer sliding windows of 200 and 300 ms, making it highly advantageous for real-time control. The observed differences highlight a fundamental trade-off: larger windows improve recognition to a certain degree but add 200–300 ms latency, while 50 ms windows deliver comparable accuracy with near-instantaneous response. Shorter windows tend to slightly increase misclassifications near transitional states due to limited

temporal context, yet dense sensor configurations may mitigate this effect through complementary spatial cues.

In practice, the model can perform immediate HAR upon receiving the first data window and continuously updates its predictions whenever a new 50 ms window is provided. From a deployment standpoint, ultra-short windows reduce computational load and energy consumption, enabling efficient on-device inference. With 50 ms windows and 50% overlap, HAR updates occur every 25 ms, providing rapid, continuous activity estimation suitable for real-time exoskeleton control. It enables the control system to switch modes rapidly in response to user activity changes with minimal delay. This balance of accuracy, responsiveness, and efficiency underscores the practical advantages of short-window, sensor-optimized configurations for embedded assistive systems.

When comparing the HAR accuracy between IMU and goniometer, we observed that IMU performs better, which is consistent with the findings of Camargo et al. [10]. Our analysis indicates that IMUs, when using only two sensors on one side of the human lower extremities, can achieve over 95% HAR accuracy with a short sliding window of 50 ms. This performance surpasses that of goniometers, particularly in differentiating various types of walking.

In terms of lower limb joint angles brought by IK data, the HAR accuracy provided by the three joints of a single (unilateral) side is slightly lower than that of three IMUs on the same side, by approximately 2.5% However, when using bilateral joint angles provided by IK, the HAR accuracy is higher than that of unilateral IMUs. It is evident from comparing Fig. 5 and 6 that when the number of joints used is reduced, the HAR accuracy using bilateral joints significantly exceeds that of unilateral joints. In addition, when using the same number of joints, the performance of using a single joint bilaterally is roughly equivalent to using two joints unilaterally. Moreover, tests with goniometers and IK show that derived joint angular velocities greatly improve HAR performance, especially with fewer joints.

The primary distinctions between our results and the work of [4] lie not only in the complexity and diversity of the activity environments captured in our utilized dataset, which include steps of various heights and ramps with different inclines, but also in our methodological advancements. Our study involves a larger number of subjects, thereby introducing greater variability and enhancing the robustness of our findings. Importantly, our approach achieved similar performance (within a 1% margin) in activity recognition using shorter sliding windows (50 vs. 200 and 300 ms), without the need for tuning the NN hyperparameters. In addition, by exclusively utilizing time-series data in LSTM, CNN-LSTM, and MLP models, our methodology simplifies the data input process, significantly reducing the computational overhead typically associated with feature extraction as used in [4]. This end-to-end learning approach highlights our study's contribution to improving the accuracy and feasibility of HAR systems in dynamic and complex environments.

Compared to [10] and [22], which used the same dataset, our study emphasizes the performance evaluation of single-type sensors (with the exception of adding a single IMU in some cases) in HAR and the impact of different joints and joint combinations under single-sensor conditions on HAR outcomes. Our results indicate that configurations relying solely on a single joint and one sensor type generally exhibit lower accuracy compared to multijoint or multimodal setups. For instance, single-joint goniometer data achieved only 50–65% accuracy [see Fig. 4(b)], whereas configurations combining joint angles with angular velocities or IMUs (e.g., hip IK + thigh IMU) reached 91–96% accuracy [see Fig. 5(b) and Fig. 8(b)]. This suggests that while single-sensor systems are inherently limited in capturing kinematic information, integrating derived angular velocity or an additional type of sensors can enhance their performance. Notably, certain single-joint configurations (e.g., hip joint angles with angular velocities or a thigh IMU) achieved accuracy levels comparable to multijoint setups, highlighting the potential for minimalist designs in specific cases. Furthermore, our findings show that the performance differences between any two joint combinations under the same sensor setup are minimal (up to approximately 2.5% between "hip+knee" and "ankle+knee" using goniometers with only joint angle data), suggesting greater flexibility in data collection for lower limb exoskeleton development. Similar to [22], we achieved comparable and superior HAR performance with setups involving a thigh IMU combined with hip joint angles and other sensors/joints. These results underscore the importance of multisensor and multijoint approaches in enhancing HAR accuracy and reliability.

The inclusion of the thigh IMUs led to improved classification accuracy for certain activities, as reflected in the confusion matrices, particularly when only a single joint type was used bilaterally (e.g., both hips or both knees etc.) Small differences in ramp inclines result in subtle kinematic variations in joint angles during incline adaptation. While sagittal joint angles alone may not be sufficiently sensitive to capture these small slope changes, combining these angles with thigh IMU data (comprising both acceleration and angular velocity) can enhance differentiation, as velocity and acceleration provide derivative information that amplifies sensitivity to minor changes in walking inclination. In practical applications, additional data sources, such as ground reaction forces measured via pressure insoles, could further improve such classification and serve as complementary inputs for HAR [10]. Moreover, improvements in model architecture could enhance HAR performance. For example, integrating an attention mechanism within the LSTM layer could enable more selective extraction of motion patterns, thereby reducing misclassification and improving overall robustness [56].

Although the configuration using all three bilateral joint angles and a thigh IMU achieved the highest performance, acquiring full joint kinematics may not be practical for many lower-limb exoskeletons. For example, adding ankle sensors to a hip or hip–knee exoskeleton introduces unnecessary complexity. In contrast, using bilateral hip and knee angles with a thigh IMU provides nearly equivalent HAR accuracy ($<$1% difference on average) while being more feasible for implementation. In the work by Molinaro et al. [7], both real-time joint angle measurements through encoders and acceleration data from IMUs were used for exoskeleton adaptive control. We believe a similar setup with encoders and IMUs can be used for HAR with our trained models for real-world deployment. In practice, if high-quality IMUs that can accurately measure joint angle and angular velocity are employed, they can potentially replace the joint encoders for HAR in real-world setup.

Despite the encouraging results obtained in this study, there are several limitations that should be noted. First, according to the description of the dataset [9], all subjects were young, healthy adults. During the experiments, the dataset authors collected IMU and goniometer data only from the right lower limb rather than from both sides. Moreover, the dataset did not include trials involving exoskeleton use. Second, this study focused exclusively on six steady-state locomotion modes and did not include

transitional activities (e.g., stand-to-walk, walk-to-stair ascent, or walk-to-ramp ascent). Third, although we achieved a 50 ms sliding window for near real-time classification in offline testing, the latency and computational loads in an actual exoskeleton controller were not validated on an integrated hardware system that includes both sensors and the HAR module. Deploying our models on such a system and measuring true closed-loop performance will be an important next step. Finally, the dataset employed in this study does not include trials involving subjects wearing active lower-limb exoskeletons, which may affect HAR performance when applied in such contexts. However, prior research has shown that healthy users often adopt gait patterns that more closely resemble their natural unassisted walking after extensive training [57] or when adaptive control strategies are employed [58]. These findings suggest that the influence of exoskeleton assistance on gait patterns, and consequently on HAR performance, may be mitigated under appropriate adaptation or control conditions. However, for practical HAR model deployment with exoskeleton usage, incorporating a small amount of exoskeleton-specific data for fine-tuning existing models represents a promising approach to further improve the model.

Future work could include evaluating the robustness of the trained models under practical conditions, with particular emphasis on clinical populations whose signal characteristics may deviate from the training data. Systematic hyperparameter tuning is also a valuable avenue for exploration, as it can materially affect final performance [59] and may reveal subtle distinctions in how different deep learning models respond to different sensor configurations and data variability. Furthermore, incorporating transitional activities into training may enhance model generalization and support smoother mode transitions within adaptive control frameworks. It would also be valuable to verify whether the ultra-short sliding window maintains the practical and energy-efficient advantages observed in this study across broader contexts. Finally, future investigations should explore closed-loop, on-device deployment on lower-limb exoskeletons, integrating real-time personalized data and online fine-tuning to reduce domain shift and achieve reliable, low-latency performance. Collectively, these directions may help translate high offline accuracy into robust, clinically meaningful performance in real-world assistive devices.

While emerging task-invariant exoskeleton controllers show promise in recent studies [5], [60], [61], they are often limited to a specific group of activities or do not demonstrate a definitive advantage in reducing metabolic costs activity-specific controllers [7]. Therefore, HAR remains highly beneficial, as it supports personalized, activity-specific controller adjustment through continuous learning, enabling seamless switching between control strategies across various activities. Furthermore, HAR facilitates the monitoring of activity-specific metrics like walking duration, step count, or number of stairs climbed etc., thus enhancing health monitoring and personalized user feedback.

VI. Conclusion

This study provides a comprehensive evaluation of sensor and data selection strategies for real-time HAR using DNN models. By analyzing a diverse dataset consisting of IMUs, goniometers, and IK derived joint angles, we highlight the critical role of sensor type, quantity, and placement on HAR accuracy. Our results demonstrate that accuracy improves with the number of joints used, and that the addition of an IMU provides the greatest benefit when only a single joint is available. Furthermore, these results consistently surpassed comparable configurations relying solely on unilateral IMUs or goniometers. Importantly, incorporating derived angular velocity information significantly enhanced accuracy, particularly evident when using single-joint bilateral IK data, underscoring the value of using derived data instead of extra sensors in HAR. Our results also reveal no significant performance differences among the three DNN models, indicating flexibility in model selection. All DNN models operate short window (50 ms) time-series data without the need for feature extraction, thus simplifying the data input process while reducing computational overhead and facilitating high frequency HAR suitable for edge computing devices such as the Jetson Xavier. Overall, our research highlights critical design considerations for implementing efficient and real-time HAR systems balancing recognition accuracy, response latency, system complexity, and power efficiency in assistive and mobile health applications.

Data and Code Availability

The code for data preprocessing, the deep learning models (MLP, LSTM, and CNN-LSTM), and inference of trained models on an edge device (Jetson Xavier) is available at https://github.com/NJITBioDynamics/HAR_DL.

References

[1] S. Luo, G. Androwis, S. Adamovich, E. Nunez, H. Su, and X. Zhou, "Robust walking control of a lower limb rehabilitation exoskeleton coupled with a musculoskeletal model via deep reinforcement learning," *J. Neuroengineering Rehabil.*, vol. 20, no. 1, p. 34, 2023.

[2] S. Luo, G. Androwis, S. Adamovich, H. Su, E. Nunez, and X. Zhou, "Reinforcement learning and control of a lower extremity exoskeleton for squat assistance," *Front. Robot. AI*, vol. 8, 2021, Art. no. 702845.

[3] K. A. Witte, P. Fiers, A. L. Sheets-Singer, and S. H. Collins, "Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance," *Sci. Robot.*, vol. 5, no. 40, 2020, Art. no. eaay9108.

[4] S. Yu et al., "Artificial neural network-based activities classification, gait phase estimation, and prediction," *Ann. Biomed. Eng.*, pp. 1–14, 2023.

[5] S. Luo et al., "Experiment-free exoskeleton assistance via learning in simulation," *Nature*, vol. 630, no. 8016, pp. 353–359, 2024.

[6] W. Huo, S. Mohammed, Y. Amirat, and K. Kong, "Fast gait mode detection and assistive torque control of an exoskeletal robotic orthosis for walking assistance," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1035–1052, Aug. 2018.

[7] D. D. Molinaro, K. L. Scherpereel, E. B. Schonhaut, G. Evangelopoulos, M. K. Shepherd, and A. J. Young, "Task-agnostic exoskeleton control via biological joint moment estimation," *Nature*, vol. 635, no. 8038, pp. 337–344, 2024.

[8] M. Pesenti, G. Invernizzi, J. Mazzella, M. Bocciolone, A. Pedrocchi, and M. Gandolla, "Imu-based human activity recognition and payload classification for low-back exoskeletons," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 1184.

[9] J. Camargo, A. Ramanathan, W. Flanagan, and A. Young, "A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions," *J. Biomech.*, vol. 119, 2021, Art. no. 110320.

[10] J. Camargo, W. Flanagan, N. Csomay-Shanklin, B. Kanwar, and A. Young, "A machine learning strategy for locomotion classification and parameter estimation using fusion of wearable sensors," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 5, pp. 1569–1578, May 2021.

[11] G. Du, J. Zeng, C. Gong, and E. Zheng, "Locomotion mode recognition with inertial signals for hip joint exoskeleton," *Appl. Bionics Biomech.*, vol. 2021, pp. 1–11, 2021.

[12] N. Bourbakis and A. Angeleas, "A synergistic formal-statistical model for recognizing complex human activities," *IEEE Trans. Human-Mach. Syst.*, vol. 54, no. 3, pp. 229–237, Jun. 2024.

[13] G. Lin et al., "Human activity recognition using smartphones with WiFi signals," *IEEE Trans. Human-Mach. Syst.*, vol. 53, no. 1, pp. 142–153, 2023.

[14] X. Chen, Y. Zou, C. Li, and W. Xiao, "A deep learning based lightweight human activity recognition system using reconstructed WiFi CSI," *IEEE Trans. Human-Mach. Syst.*, vol. 54, no. 1, pp. 68–78, Feb. 2024.

[15] C. Jobanputra, J. Bavishi, and N. Doshi, "Human activity recognition: A survey," *Procedia Computer Science*, vol. 155, pp. 698–703, 2019, the 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2019), The 14th Int. Conf. Future Netw. Commun. (FNC-2019), The 9th International Conference on Sustainable Energy Information Technology.

[16] W. Qi, H. Su, and A. Aliverti, "A smartphone-based adaptive recognition and real-time monitoring system for human activities," *IEEE Trans. Human-Mach. Syst.*, vol. 50, no. 5, pp. 414–423, 2020.

[17] D. Anguita et al., "A public domain dataset for human activity recognition using smartphones," in *The Eur. Symp. Artif. Neural Netw.*, vol. 3, 2013, p. 3.

[18] P. Zappi et al., "Activity recognition from on-body sensors: Accuracy-power trade-off by dynamic sensor selection," in *Wireless Sensor Networks: 5th Eur. Conf.*, EWSN 2008, Bologna, Italy, Jan. 30-February 1, 2008, Proceedings. Springer, 2008, pp. 17–33.

[19] H. Yu, S. Cang, and Y. Wang, "A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems," in *Proc. 10th Int. Conf. Software, Knowledge, Inf. Manage. Appl.*, 2016, pp. 250–257.

[20] H. Haresamudram, D. V. Anderson, and T. Plötz, "On the role of features in human activity recognition," in *Proc. ACM Int. Symp. Wearable Comput.*, 2019, pp. 78–88.

[21] U. H. Lee, J. Bi, R. Patel, D. Fouhey, and E. Rouse, "Image transformation and cnns: A strategy for encoding human locomotor intent for autonomous wearable robots," *IEEE Robot. Automat. Lett.*, vol. 5, no. 4, pp. 5440–5447, Oct. 2020.

[22] I. Kang, D. D. Molinaro, G. Choi, J. Camargo, and A. J. Young, "Subject-independent continuous locomotion mode classification for robotic hip exoskeleton applications," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 10, pp. 3234–3242, Oct. 2022.

[23] L. Bergmann et al., "Lower limb exoskeleton with compliant actuators: Design, modeling, and human torque estimation," *IEEE/ASME Trans. Mechatron.*, vol. 28, no. 2, pp. 758–769, 2023.

[24] J. Wang et al., "A selection framework of sensor combination feature subset for human motion phase segmentation," *Inf. Fusion*, vol. 70, pp. 1–11, 2021.

[25] Z. Chen, Q. Zhu, Y. C. Soh, and L. Zhang, "Robust human activity recognition using smartphone sensors via ct-pca and online svm," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3070–3080, Dec. 2017.

[26] J. Li et al., "Feature selection: A data perspective," *ACM Comput. Surv.*, vol. 50, no. 6, Dec. 2017.

[27] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, 2020, Art. no. 132306.

[28] R. Jain, V. B. Semwal, and P. Kaushik, "Deep ensemble learning approach for lower extremity activities recognition using wearable sensors," *Expert Syst.*, vol. 39, no. 6, 2022, Art. no. e12743.

[29] K. Xia, J. Huang, and H. Wang, "Lstm-cnn architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56855–56866, 2020.

[30] Y. Chen, K. Zhong, J. Zhang, Q. Sun, and X. Zhao, "Lstm networks for mobile human activity recognition," in *Proc. Int. Conf. Artif. intelligence: Technol. Appl.*. Atlantis Press, 2016, pp. 50–53.

[31] M. Ronald, A. Poulose, and D. S. Han, "isplinception: An inception-resnet deep learning architecture for human activity recognition," *IEEE Access*, vol. 9, pp. 68985–69001, 2021.

[32] I. Dirgová Luptáková, M. Kubovčık, and J. Pospıćhal, "Wearable sensor-based human activity recognition with transformer model," *Sensors*, vol. 22, no. 5, 2022, Art. no. 1911.

[33] C. Xia and Y. Sugiura, "Optimizing sensor position with virtual sensors in human activity recognition system design," *Sensors*, vol. 21, no. 20, 2021, Art. no. 6893.

[34] Z. Zheng, H. Ma, W. Yan, H. Liu, and Z. Yang, "Training data selection and optimal sensor placement for deep-learning-based sparse inertial sensor human posture reconstruction," *Entropy*, vol. 23, no. 5, 2021.

[35] C. Caramia et al., "Imu-based classification of parkinson's disease from gait: A sensitivity analysis on sensor location and feature selection," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 6, pp. 1765–1774, 2018.

[36] G. L. S and A. Poulose, "From wrist to ankle: Understanding imu sensor placement in human activity recognition," in *Proc. Emerg. Technol. Intell. Syst.*, 2025, pp. 1–6.

[37] V. X. Rahn, L. Zhou, E. Klieme, and B. Arnrich, "Optimal sensor placement for human activity recognition with a minimal smartphone-imu setup," in *Int. Conf. Sensor Netw.*, 2021.

[38] P. N. Müller, A. J. Müller, P. Achenbach, and S. Göbel, "Imu-based fitness activity recognition using cnns for time series classification," *Sensors*, vol. 24, no. 3, 2024.

[39] M. P. Kadaba, H. Ramakrishnan, and M. Wootten, "Measurement of lower extremity kinematics during level walking," *J. orthopaedic Res.*, vol. 8, no. 3, pp. 383–392, 1990.

[40] M. Boyer, L. Bouyer, J.-S. Roy, and A. Campeau-Lecours, "Reducing noise, artifacts and interference in single-channel emg signals: A review," *Sensors*, vol. 23, no. 6, 2023, Art. no. 2927.

[41] A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Appl. Soft Comput.*, vol. 62, pp. 915–922, 2018.

[42] Y. Li et al., "A single smartwatch-based segmentation approach in human activity recognition," *Pervasive Mobile Comput.*, vol. 83, 2022, Art. no. 101600.

[43] C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, "Innohar: A deep neural network for complex human activity recognition," *Ieee Access*, vol. 7, pp. 9893–9902, 2019.

[44] R. Raj and A. Kos, "An improved human activity recognition technique based on convolutional neural network," *Sci. Rep.*, vol. 13, no. 1, 2023, Art. no. 22581.

[45] N. Sikder, M. S. Chowdhury, A. S. M. Arif, and A.-A. Nahid, "Human activity recognition using multichannel convolutional neural network," in *Proc. 5th Int. Conf. Adv. Elect. Eng.*, 2019, pp. 560–565.

[46] S. A. Mohamed and U. Martinez-Hernandez, "A light-weight artificial neural network for recognition of activities of daily living," *Sensors*, vol. 23, no. 13, 2023, Art. no. 5854.

[47] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "A deep learning-based model for human activity recognition using biosensors embedded into a smart knee bandage," *Procedia Comput. Sci.*, vol. 214, pp. 621–627, 2022.

[48] D. Wagner, K. Kalischewski, J. Velten, and A. Kummert, "Activity recognition using inertial sensors and a 2-D convolutional neural network," in *Proc. 10th Int. Workshop Multidimensional (nD) Syst.*, IEEE, 2017, pp. 1–6.

[49] R. R. Drumond, B. A. D. Marques, C. N. Vasconcelos, and E. Clua, "An lstm recurrent network for motion classification from sparse data," in *Proc. 13th Int. joint Conf. Comput. Vis.*, imaging and computer graphics theory and applications, vol. 1, 2018, pp. 215–22.

[50] W.-H. Chen et al., "Determining motions with an IMU during level walking and slope and stair walking," *J. sports Sci.*, vol. 38, no. 1, pp. 62–69, 2020.

[51] F. Sherratt, A. Plummer, and P. Iravani, "Understanding LSTM network behaviour of IMU-based locomotion mode recognition for applications in prostheses and wearables," *Sensors*, vol. 21, no. 4, 2021, Art. no. 1264.

[52] H. Lu, L. R. Schomaker, and R. Carloni, "IMU-based deep neural networks for locomotor intention prediction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*. IEEE, 2020, pp. 4134–4139.

[53] K. Suwannarat and W. Kurdthongmee, "Optimization of deep neural network-based human activity recognition for a wearable device," *Heliyon*, vol. 7, no. 8, 2021.

[54] J. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto, and X. Parra, "Human activity recognition using smartphones," *UCI Mach. Learn. Repository*, 2012, doi: 10.24432/C54S4K.

[55] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, Mar. 2011. [Online]. Available: https://doi.org/10.1145/1964897.1964918

[56] K. Muhammad et al., "Human action recognition using attention based LSTM network with dilated CNN features," *Future Gener. Comput. Syst.*, vol. 125, pp. 820–830, 2021.

[57] A. Lakmazaheri and S. H. Collins, "Characterizing expert exoskeleton-assisted gait: Insights to accelerate ankle exoskeleton mastery," *IEEE Trans. Med. Robot. Bionics*, pp. 1–1, 2025.

[58] M. Shushtari, J. Foellmer, and A. Arami, "Human–exoskeleton interaction portrait," *J. NeuroEngineering Rehabil.*, vol. 21, no. 1, 2024, Art. no. 152.

[59] J. Suto and S. Oniga, "Efficiency investigation of artificial neural networks in human activity recognition," *J. Ambient Intell. Humanized Comput.*, vol. 9, pp. 1049–1060, 2018.

[60] D. D. Molinaro, I. Kang, and A. J. Young, "Estimating human joint moments unifies exoskeleton control, reducing user effort," *Sci. Robot.*, vol. 9, no. 88, 2024, Art. no. eadi8852.

[61] K. Walters, G. C. Thomas, J. Lin, and R. D. Gregg, "An energetic approach to task-invariant ankle exoskeleton control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*. IEEE, 2023, pp. 6082–6089.