Open-Source Virtual IMU Sensor Platform for Developing a Customized Wearable Human Activity Recognition System

Chengshuo Xia, *Member, IEEE*, Xinyi Chen, Lingtao Huang, Tian Min, and Daxing Zhang

***Abstract*— Building a wearable human activity recognition (HAR) system follows complicated steps, making the wearable HAR system development time-consuming, cumbersome, and error-prone. One of the critical challenges is its dataset collection, which typically requires a group of people to perform specified motions for an extensive period to capture data. Recently, cross-modal virtual inertial measurement unit (IMU) data generation from 3-D avatars has been recognized as a key solution to address the costly and laborious sensor dataset collection for HAR system establishment. However, the utilization of virtual IMU data still faces the challenge of tedious steps, including expensive computing resources, model training, and domain transfer, among others. This article presents a novel platform employing the input of 3-D avatar motion sequences. Combined with the virtual IMU data generation mechanism, the motion modification-based virtual IMU data augmentation approach is designed to generate the virtual IMU training dataset. The platform supports various sensor placement selections, data processing, model training, fine-tuning, and deployment functionalities. The proposed platform integrates the whole process of developing a wearable HAR system. Due to the fewer requirements in the real world, it could enable the wearable HAR system building to become low-cost, customized, and contribute to HAR system prototyping and related application scenarios.**

***Index Terms*— 3-D avatar, human activity recognition (HAR), platform, virtual inertial measurement unit (IMU), wearables.**

I. Introduction

Wearable human activity recognition (HAR) system has been drawing close attention in building a natural and seamless interface between the human and computing system, such as in entertainment, safety monitoring, rehabilitation, and fitness tracking systems [1], [2], [3]. The mainstream HAR system utilized the inertial measurement unit (IMU) to capture the human motion data and recognize the activities via a machine learning model. The corresponding technique routine has been widely applied to enormous end-user electronics, including smartphones, smartwatches, smart glasses, and so on [4], [5], [6], [7].

Though traditional wearable HAR systems have been realized to a certain extent with successful applications and have presented huge benefits in assisting users' daily lives, demands from users are moving toward greater variety. For example, as various users may have different exercise habits and body conditions, the wearable HAR system of fitness tracking is required to recognize motions based on the user's distinct needs. Therefore, the end-user's requirements for such sensor-based machine learning systems are becoming increasingly diverse, and the system needs to recognize objects or motions that can be customized and personalized. Generally, the traditional wearable HAR system development follows a typical technical chain from dataset collection, data processing, model training, and testing to output an ultimate model in practice [8], [9]. The process involves several steps, such as individual recruitment, data collection, annotation, and model development, resulting in cumbersome, inefficient, and error-prone characteristics. It makes the development face several difficulties for the end users in establishing a customized HAR system. In particular, collecting the dataset has been recognized as the most challenging part [10]. As a time-consuming process, data collection is often laborious and has high resource demands. End users usually fail to create their own dataset for the wearable HAR system building and utilizing to satisfy their personalized needs.

Different from the vision-based system, public wearable IMU-based datasets have relatively small sizes and numbers of motion types, making it hard to apply the model in more generalized scenarios. Existing solutions either mitigate this acute problem from the perspective of recognition algorithms or HAR system development tooling techniques. Specifically, zero-shot or few-shot learning attempts to apply already pre-trained models to additional motion recognition via the knowledge transferring approach [11], [12]. However, the ability of the IMU data representation still limits the ability of such methods, and the existing research results have merely proved the effectiveness of a few simple types of motions. In addition, to lower the development threshold for the wearable HAR system, the integrated tooling or engineering platform techniques would contribute to a fast and simple development process [13], [14]. Thus, researchers have also proposed development tools for easy development, including

Received 25 August 2024; revised 8 December 2024; accepted 20 December 2024. Date of publication 5 March 2025; date of current version 20 March 2025. This work was supported in part by China Postdoctoral Science Foundation under Grant 2024M762548, in part by the Fundamental Research Funds for the Central Universities under Grant XJSJ23109, and in part by the High-Level Innovation Institute Project of Guangdong Province under Grant 2021B0909050008. The Associate Editor coordinating the review process was Dr. Sergio Rapuano. *(Corresponding author: Chengshuo Xia.)*

Chengshuo Xia is with the Advanced Manufacturing Technology Innovation Center, Guangzhou Institute of Technology, Xidian University, Guangzhou 510555, China (e-mail: xiachengshuo@xidian.edu.cn).

Xinyi Chen, Lingtao Huang, and Daxing Zhang are with Guangzhou Institute of Technology, Xidian University, Guangzhou 510555, China.

Tian Min is with the Graduate School of Science and Technology, Keio University, Minato 108-0073, Japan.

Digital Object Identifier 10.1109/TIM.2025.3548063

data annotation software, model training codes, and visualization tools [16], [17]. Nevertheless, these methods still adopt the traditional HAR system development chain. It is difficult to bypass the laborious, costly, and bulky characteristics of the real data collection process.

On the contrary, cross-modal sensor data generation has received much attention recently [18]. The generation pipeline from the video to 3-D avatar motion to virtual IMU data demonstrates new capabilities in data collection for the wearable HAR system [19], [20]. This approach enables the introduction of large video datasets into the IMU dataset to address the data scarcity issue and improve the wearable HAR system's performance. The virtual IMU data has been utilized in sign language recognition, multimodal HAR system development, and text-based IMU synthesis [21], [22], [23]. However, existing virtual IMU applications aim to improve the generalizability characteristics of HAR by expanding the size of existing datasets, showing the constrained applications. How to further extend the application of virtual IMU to improve more HAR scenarios is not well explored, especially for utilizing the low-cost and flexible characteristics of the virtual IMU data generation process.

Therefore, in this article, we proposed an open-source platform using virtual IMU data to develop a customized HAR system on the user end (Fig. 1). It supports the users in indicating personalized recognized motions and includes the dataset generation and model training process to build a wearable HAR system. Compared with the traditional technique chain, development based on our platform is realized merely in the software environment. Users can select the sensor locations and numbers with great flexibility, and recruiting real user groups for the collection of large real datasets is eliminated. At the user end, a few-shot real data are collected to fine-tune the trained model as a calibration process to establish the corresponding customized wearable HAR system. Section II introduced the related work, and Section III presented the details of platform building. We also conducted several experiments to demonstrate the feasibility of using the platform to design a practical wearable HAR system, as shown in Section IV.

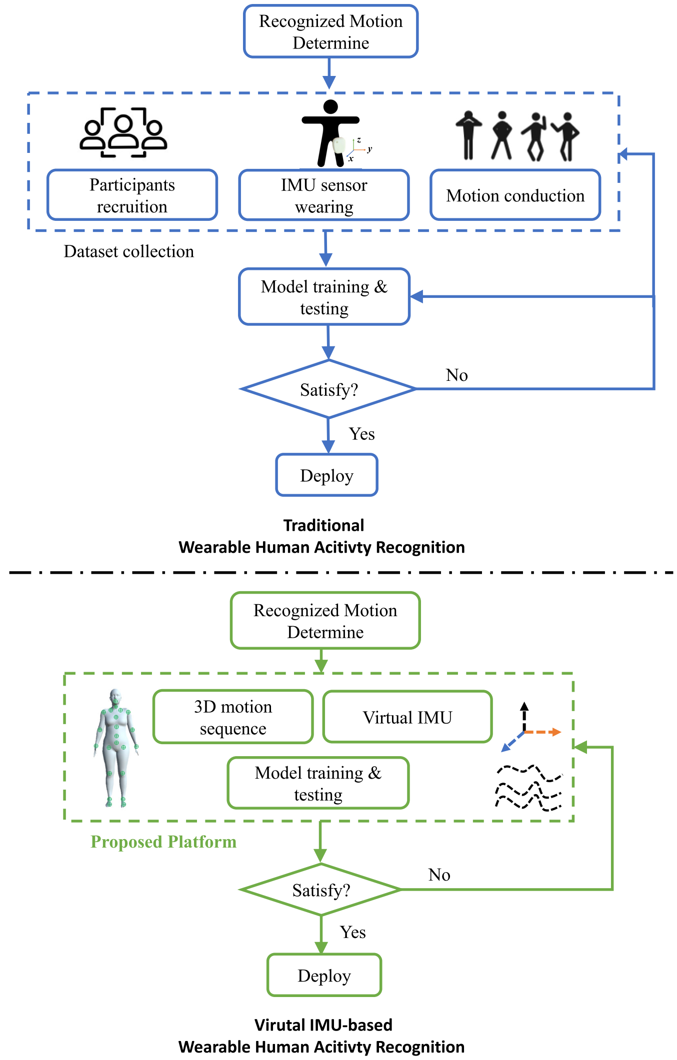


Fig. 1. Comparison between the traditional HAR system development process and the proposed platform for customized HAR system development using virtual IMU data.

II. Related Work

A. Generalized and Customized HAR System

IMU-based HAR, as the pillar of natural and ubiquitous activity recognition systems, has already been applied to numerous commercial-of-the-shelf electronics, such as smart watches, smart rings, and glasses [4], [5], [6], [7]. Commonly, the developed wearable HAR system needs to overcome the challenge from the changing position of attachment, the difference in users' motion habits, and individual characteristics [9], [24]. Thus, it requires the wearable HAR system to maintain high generalizability. At the data generation level, the synthesis sensor data can effectively increase the training dataset, raising the performance of the machine learning model. For example, Zilelioglu et al. [25] presented a semi-supervised generative adversarial network (GAN) to solve the problem of the scarcity of annotated IMU data in the wearable HAR system. Um et al. [26] used the signals of Parkinson's patients and proposed to use several data modification approaches for the data augmentation, including time-wrapping, rotation, and cropping. The modified signal mixed with the original signal could enhance the trained CNN model for Parkinson's disease monitoring. In addition, at the data representation level, Su et al. [24] designed a disentangled method for IMU data using the GAN model to address the intra-class variability issue and disentangled the dataset of the wearable HAR system into intra-class invariant features and redundant features, thus improving the generalization ability of the system. Xia et al. [27] introduced a multiple-level domain adaptive learning model with information theory to align the distribution of both virtual IMU and real IMU data representation. From the training approach aspect, incremental learning has brought new solutions and gained attention to help the trained model to adapt to various scenarios [28], [29], [30]. For example, Hou et al. [29] introduced a two-stage simultaneous augmentation of feature and class, which employed regularizers for model inheritance and reuse. Hu et al. [28] presented the feature incremental random forest (RF) method to improve the model's performance on new features. It involved a mutual information-based strategy to enrich the diversity of the RFs, and a feature incremental growing mechanism assisted in the accuracy of each decision tree.

On the other hand, research on customization and personalization for activity recognition has been continuously proposed, particularly in the human-computer interaction scenarios. The zero- and few-shot learning create the alternative solutions for applying a pre-trained machine learning model to the front-end applications with the increased customization factor [11], [12]. For example, Su et al. [31] showed a few-shot learning-based lip language reading in a personal smartphone via the in-front camera. Xu et al. [32] developed a wrist-worn smartwatch-based hand gesture customization approach for end users. The collected hand gesture IMU dataset was utilized to pre-train the recognition model, and few-shot learning was adopted to fine-tune the model and realize a user-centered hand gesture recognition process. Steuerlein and Mayer [33] employed the GAN model to assist the capacitive image generation in training a customized model for the tangible interaction with a touchscreen in a tablet. Therefore, enhancing the customization of HAR systems has gradually become the research focus. However, the development of existing HAR systems customization is based on real IMU datasets, which still require people to contribute many motions and data. In this article, we focus on exploiting virtual IMU and leveraging its low-cost and easy-to-generate characteristics to develop its customized HAR system.

B. Cross-Modal Virtual IMU Data Utilization

The virtual IMU data could be extracted from a 3-D avatar motion sequence in the virtual environment [36]. It follows a cross-modal pipeline that extracts the human skeleton information from 2-D human motion video and reconstructs the 3-D motion by a 3-D avatar [19]. Generating virtual IMU data provides a more convenient, low-cost, and intuitive data generation approach compared with traditional data-driven methods (e.g., GAN-based data synthesis). Virtual IMU provides additional data sources to expand existing HAR public datasets, such as IMUTube [19] and sign language recognition [21]. It therefore demonstrates the advantage of model generalizability enhancement.

One of the major bottlenecks in applying virtual IMU is the quality of virtual IMU data. The mainstream video-based methods for generating 3-D motion sequences are affected by video occlusion and quality. In addition, the position of the wearable IMU and noise issues also affect the difference between virtual IMU and real IMU data. Therefore, improving the data quality and application of virtual IMU has become the focus of many researches [18], [37]. For example, Xia and Sugiura [18] proposed a virtual spring model to augment the virtual IMU data from a few motion sequences as input. The designed spring model could utilize the physical simulation characteristic to compensate for and augment the virtual IMU data quality. Generating the virtual IMU data has been under constant scrutiny. However, the exploration of its application still needs to be expanded to augment existing public datasets and enhance the generalization performance of recognition models. In this article, we employed a method of augmenting virtual IMU data by modifying the 3-D avatar motion sequence and integrating the function within the HAR system development platform.

C. HAR Development Tools and Platforms

To better facilitate the development of the wearable HAR system, researchers have contributed to the work of tooling techniques to help reduce the difficulty and tediousness of the development process. Haladjian [17] presented an integrated development environment for wearable HAR system design. They utilized the MATLAB script for packaging-related recognition algorithms and test functions aimed at lowering the barrier in building the IMU-based HAR system. Similarly, Schipor and Vatavu [38] developed a software tool to help the experiment-centered design in wearable gesture recognition. It can support gestures designed by different equipment via HTTP and WebSocket communication. Karolus et al. [39] introduced a toolkit to facilitate the electromyography-based human gesture recognition system for human–computer interaction. The designed toolkit could connect to the electrodes, define the gesture, and be calibrated by users. After collecting the data, the toolkit also supported training the model and classifying the gesture in real-time. Ding et al. [40] presented a free-weight exercise monitoring platform. The system utilized the radio frequency identification (RFID) tag on the dumbbells and the Doppler shift profile of the reflected backscatter signals to recognize the difference motion exercise. It integrates signal preprocessing, data segmentation, recognition, activity post-assessment, and other crucial steps.

In summary, the tools facilitating the wearable HAR system development normally have the basic steps of data collection and processing. The machine learning-based platform needs the steps for model training and testing. Moreover, the feedback on model accuracy is vital for the developers to adjust the system parameters and configuration during the trial-and-error process. However, the current platform usually operates under the premise of the existing off-the-shelf datasets and provides the developers with few functions to adjust, thus less flexibility. Once the dataset is flawed or unsatisfactory, the developer has to recruit personnel and collect datasets again, which is time-consuming and labor-intensive. This article presented an alternative solution relying on the ease of accessing the virtual IMU data rather than traditional real IMU data collection. The development of the virtual IMU data-based HAR system promises a new paradigm for wearable HAR systems prototyping and development.

III. Platform Design

A. Pipeline

Generating the virtual IMU data from the 3-D avatar motion could be a digital twin of capturing the real IMU data from the human body. Thus, our platform inputs the 3-D motion sequences of desired recognized activities as the source of virtual IMU data. The entire platform was developed based on the game engine Unity3D. Fig. 2 shows the pipeline of

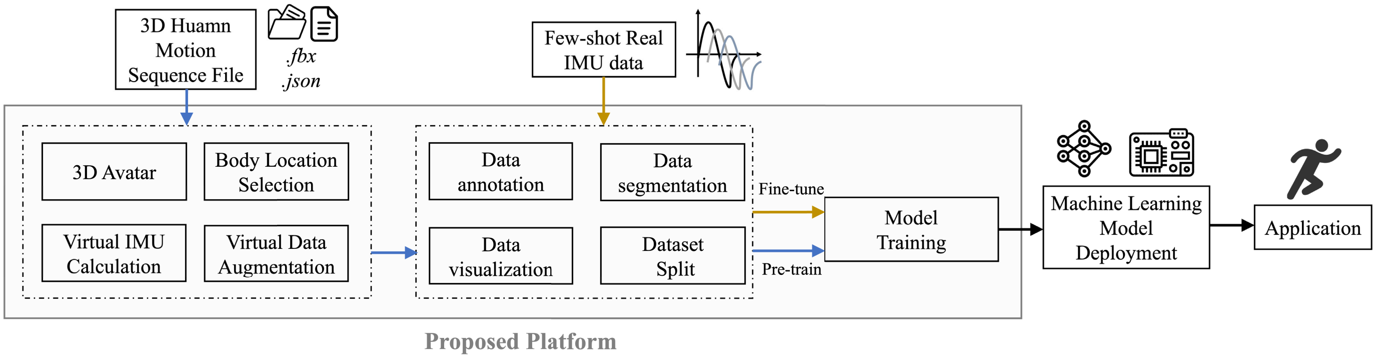


Fig. 2. Illustration of the proposed platform for customized HAR system development based on virtual IMU data.

the proposed platform. The virtual IMU data is derived from the kinematic equations [e.g., (1)]. The specific body limb is accessed to get the position $\boldsymbol{p} \in \mathbb{R}^3$ and rotation quaternion $\boldsymbol{q} \in \mathbb{R}^4$. The acceleration signal could be calculated by secondary differential [36]

$$\boldsymbol{a}(t) = \boldsymbol{q}^*(t) \otimes \frac{p_{t-\Delta t} + p_{t+\Delta t} - 2p_t}{\Delta t^2} \otimes \boldsymbol{q}(t) \tag{1}$$

where the $\boldsymbol{a}(t)$ is the calculated acceleration data as the virtual IMU acceleration data. The $\boldsymbol{q}(t)$ is the quaternion of the body limbs, and $\boldsymbol{q}^*(t)$ denotes the conjugate of the quaternions. p_t represents the position of the body limbs in a motion sequence. Δt is the sampling time interval that can control the virtual IMU data sampling rate. The $\otimes$ represents the quaternion multiplication operator.

The wearable HAR machine learning model in the proposed platform is trained based on the virtual IMU dataset. Since there are normally differences between the virtual IMU data and real IMU data (e.g., the various wearing positions and the noise of 3-D motion conversion), a transfer learning-based model fine-tuning approach is adopted to compensate for the data domain difference between the virtual IMU data and real IMU data. We allowed the user to provide a small amount of personal real IMU data to adapt the pre-trained machine learning model with virtual IMU data to recognize the personal real activity. Thus, a real IMU device interface is also designed in the platform and is able to gather a few-shot real data for model transfer.

B. 3-D Avatar Motion

The ways of generating 3-D avatar motion sequences are varied and include manual design, reconstruction using the motion capture (MoCap) device, cross-modal approaches, and the text-based generation method [41]. To ensure a general imported motion file format, the developed platform is able to support two kinds of motion files to allow the 3-D avatar to conduct the motion in the virtual environment, i.e., the *fbx* file and *json* file with the skeleton information. The *fbx* file has more complete information on the 3-D avatar's movement, including the skeletal rotation data and mesh information. It is the common output file format of the commercial off-the-shelf (COTS) MoCap device and human motion generation software.

C. Motion Modification-Based Virtual IMU Data Augmentation

Since the input 3-D motion sequence length affects the size of the virtual IMU dataset, augmentation of the virtual IMU data is necessary in customized HAR system building. In this section, we presented a data augmentation method for virtual IMU data. Referring to the situation in the real world, when an individual performs a motion, the real IMU data generated from the same motion have different amplitudes and speeds. Therefore, the proposed virtual IMU data augmentation method focused on simulating the intra-difference of IMU data caused by different users for the same motion.

In the virtual environment, the single body-limb movement of a 3-D avatar can be represented by rotation and position. To animate a human motion, rotation plays a more important role. Thus, by modifying the rotation attributes of the 3-D avatar's joints in the virtual environment, the original avatar's motion could be altered to generate more 3-D motions and obtain the augmented virtual IMU data.

To enable the generation of intra-difference for the 3-D avatar, the amplitude of the joints is modified by using multiplication factors and subsequently altered by the filters to produce more variability of the motions. The imported motion sequence $\{x_i\} = \{x_1, x_2, x_3, \ldots, x_t\}$, where t is the total number of motion frames. In this article, the motion sequence is represented by the avatar joint rotations, so $\{x_i\} \in \mathbb{R}^{j \times d}$, where j is the number of joints, and d is the dimension of the rotation representation (here we used the Euler angle and d is three). We first change the animation by modifying the amplitude of the joint motions, and a multiplication factor is set to modify the amplitude of the 3-D avatar's joint motion linearly.

We then utilize filters to modify the joint rotation data. The kernel's design of the filter can be considered a function that weighs and manipulates the input signal. Different kernels can change the characteristics of the input signal, such as frequency components, amplitude, and phase. Therefore, we consider the motion sequences as input signals and change their frequency and time domain information by designing different kernels to increase the intra-difference of the virtual IMU data. The method of modifying the motion sequence can be expressed by the following equation:

$$x_i^{\text{axis}} = \alpha\{k\} \otimes x_i^{\text{axis}} \tag{2}$$

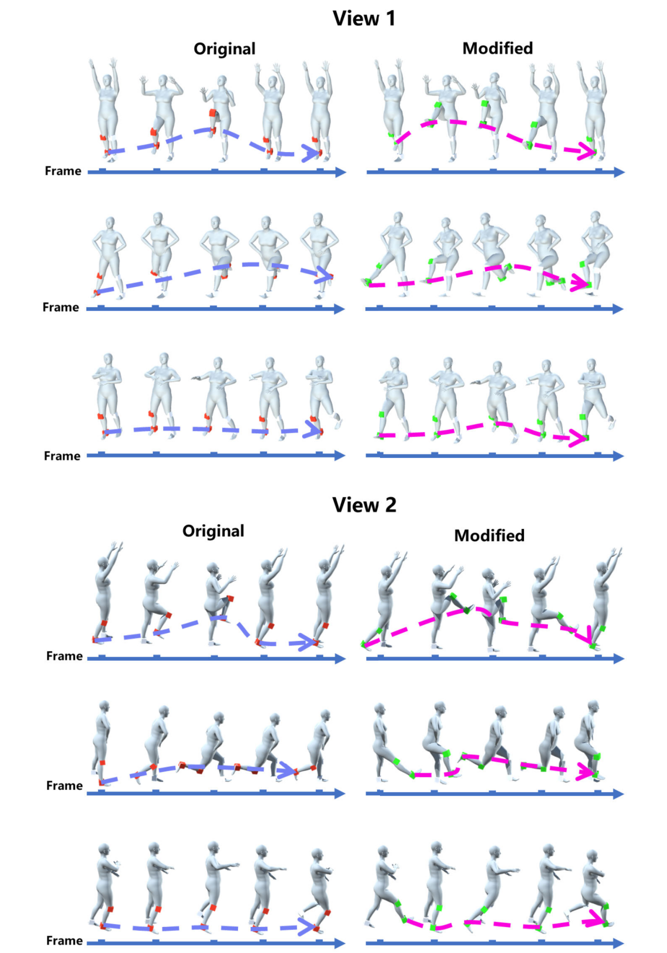


Fig. 3. Original motion versus modified motion, where the x-axis of the avatar's *right hip* has been modified. The top to bottom motions are *reverse lunge*, *high knee tap*, and *side to side*.

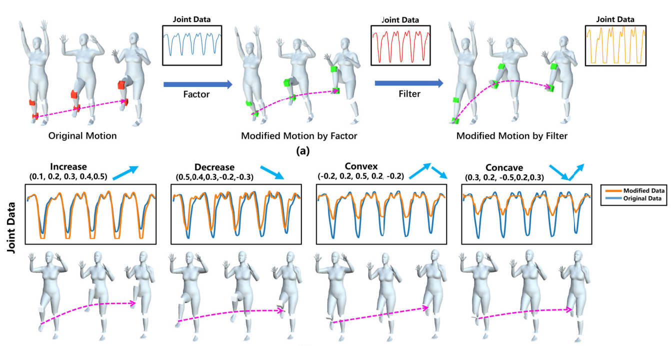


Fig. 4. Illustration of the proposed 3-D motion modification-based virtual IMU data augmentation method. (a) and (b) Modification effects on 3-D motion. (b) Skeleton information of an avatar.

where α is the amplitude change factor. We have referenced the human motion ergonomic limitation to ensure that the avatar's 3-D animation is realistic and natural [42]. The value of α is set between 0.5 and 1.6, where k is the filter kernel, and i and axis denote the joint sequence number and the joint's axis to be modified, respectively.

To verify that the filter can implement the modification of the joint data and obtain the new 3-D motions, we designed different parameter types (increase, decrease, convex, concave) for the kernel with a window size of 5. Increase means that the parameter in the window is gradually growing, decrease represents that the parameter is reducing progressively, and convex implies that the parameter is growing and then reducing. Concave means that the parameter is reducing and then increasing. These four parameter types can compose different filters, so we use these four parameter types of filters to verify whether they can complete the joint data modification. Fig. 3 shows the overview comparison between the initial and modified motions. The kernel parameters, the parameter types, and the modified joint data are shown in Fig. 4. Fig. 4 shows an example of avatar motion. The proposed method could modify the major moving limbs to produce augmented movements and virtual IMU data. Thus, due to the modification of the joint data by the filter, it could result in a noticeable change in the *right hip* motion of the 3-D avatar. In addition, we artificially set the maximum motion limits of the avatar's joints in the virtual environment so that the modified motion sequences did not show non-natural motion [37]. The example of the virtual IMU from the original motion versus the modified motion is shown in Fig. 5. The developed data augmentation method simulates the intra-class variability of movements between real people. This method is simple and efficient, modifying the limb movement of the 3-D avatar based on the position of the wearable sensors to create a greater distribution of movements, thus generating more virtual IMU data.

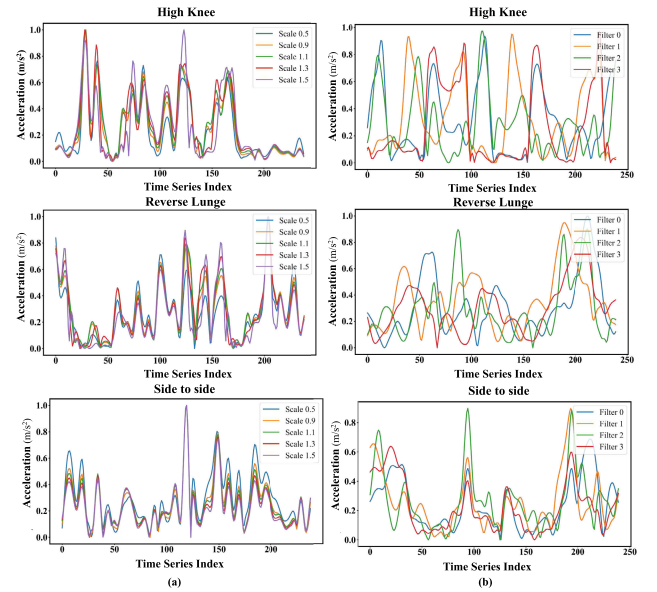


Fig. 5. Virtual IMU data generated by the proposed data augmentation method. (a) Modified by a factor with different parameters. (b) Modified by a filter with different parameters.

D. Transfer Learning-Based Model Training From Virtual IMU to Real IMU Data

The generated virtual IMU dataset was subsequently employed to train the machine learning model. The IMU signals from the virtual and real domains differ in acceleration signal amplitude, coordinate system transformation, and noise error. Therefore, we designed the model structure, which is prone to perform a domain adaption-based transfer learning

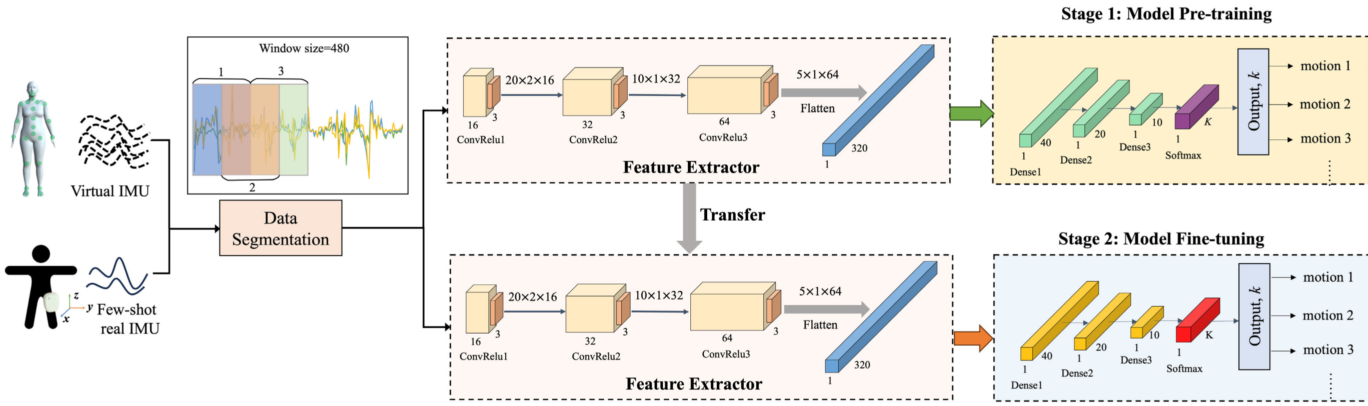


Fig. 6. Proposed CNN structure used for virtual IMU-based customized HAR system building in developed platform.

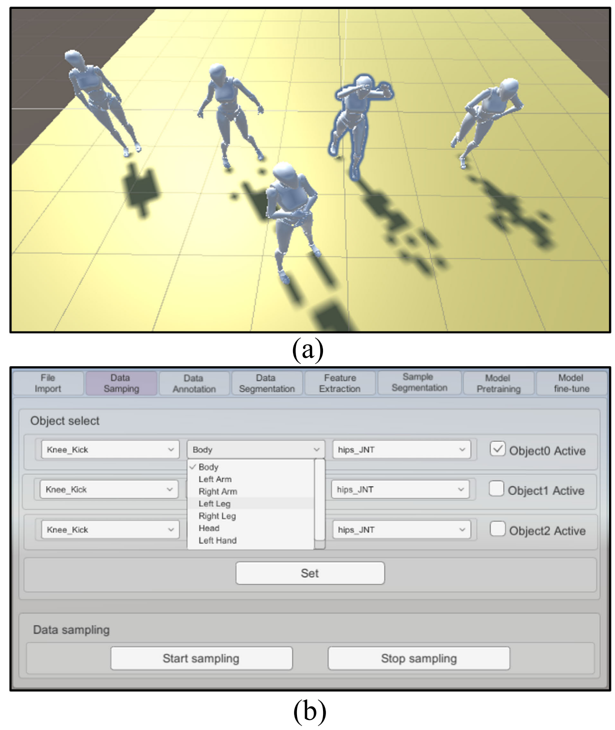
from the virtual to the real domain. Inspired by self-supervised learning, the whole transfer learning follows the pre-trained and fine-tuned process [43], [44]. The virtual IMU dataset is first used to train the model in a supervised way, and the model's partial weights are frozen, and a few layers are fine-tuned by real IMU data in a supervised way as well.

Fig. 6 introduces the structure of the model in the proposed platform. The data processing and segmentation steps generated training data for the input IMU data. In the pre-training stage, the augmented virtual IMU data is utilized to train the feature extraction layers, including three convolutional layers plus pooling layers and output with a flattened layer. It aims to ensure a feature extraction ability for input signals. Then, to bridge the gap between the virtual and real domains, three dense layers are re-trained by a few-shot real IMU data samples collected from the individuals conducting the classified motions. This step would adjust the dense layers to make the whole CNN model transfer more suitable for real IMU data recognition.

E. Functions and Interfaces

The platform integrated the representative sensor data processing functions in a typical sensor-based machine learning system-building process, including data annotation, segmentation, and dataset split for testing. Therefore, several functions and sub-interfaces were created in the platform.

1) Motion Files Import and Configuration: In the beginning, the platform supports reading uploaded motion sequence files with user-designated names. Then, the platform enters into the data configuration step. It could set the virtual IMU data sampling frequency. The user also determines the numbers and locations of the virtual IMU used (Fig. 7). The platform identifies multiple optional wearing locations based on the hierarchical structure of the 3-D avatar, including the *head, neck, right upper arm, right wrist, left upper arm, left wrist, waist, right upper leg, right lower leg, left upper leg, left lower leg, right foot, and left foot*. The tentative wearing positions are able to be connected to various daily electronics such as smart glass/earphone in the head and intelligent knee-pad in the lower leg. The motion files from the motion modification-based data augmentation method are



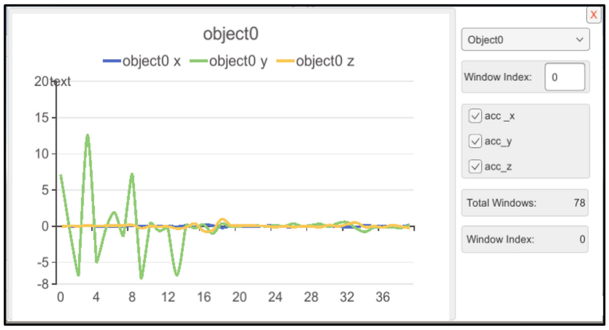
(a) (b)

Fig. 7. Interfaces for 3-D avatar motion import and data generation configuration. (a) Three-dimensional avatars. (b) Configuration of sensor locations.

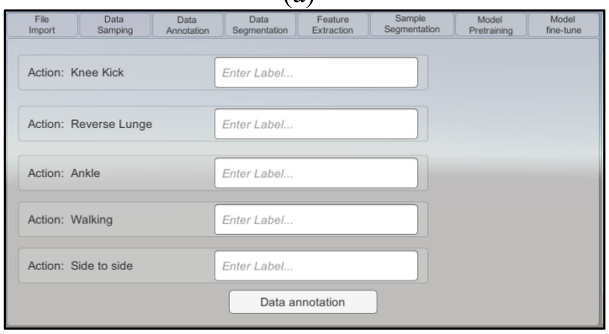
simultaneously imported into the platform to produce the virtual IMU data as well.

2) Virtual IMU Processing: After the virtual IMU data is generated, the platform allows the user to perform the annotation and segmentation. Fig. 8 presents the corresponding sub-interfaces for data labeling and splitting. The user could indicate the arbitrary length of data frames. Unlike the vision-based data sample, IMU data is abstract and meaningless. So, the platform also provides a visualization tool to assist the user in observing the segmented data frame by each axis and determining if it is necessary to adjust the length.

3) Model Structure Configuration: The CNN model configuration is also embedded into the platform to enable the user to flexibly design the required model structure, including the training epochs, optimizer, learning rate, and loss function.

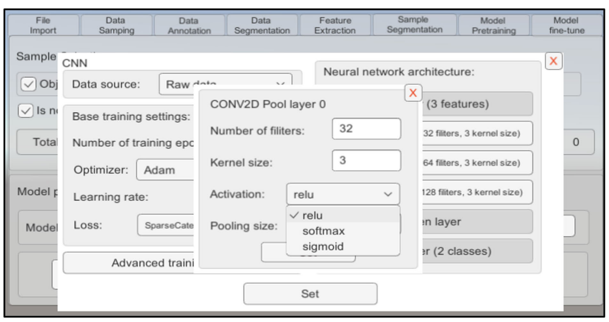


(a)

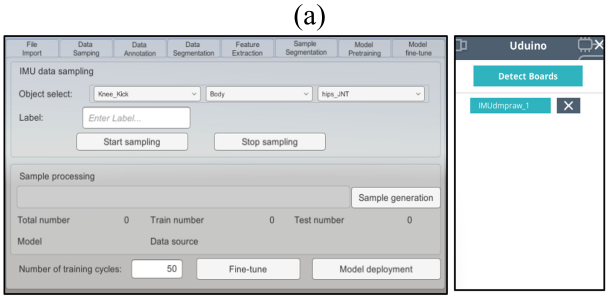


(b)

Fig. 8. Virtual IMU data visualization and annotation. (a) Virtual IMU data visualization. (b) Data annotation.



(a)



(b)

Fig. 9. CNN model structure configuration and fine-tuning with real IMU data. (a) CNN model configuration. (b) Model fine-tune and deployment.

In addition, the user is allowed to configure the structure as well. For example, determining how many convolutional layers and dense layers are used and designing the parameters of each layer. The related sub-interfaces are shown in Fig. 9.

4) Testing and Fine-Tuning: After the model configuration, the platform could split the virtual IMU dataset for the model's performance test at first. It will provide feedback on the preliminary evaluation results to help the user modify the related system setting accordingly. Due to the requirement of domain transfer, Fig. 9 also presents the model's fine-tuning process.

Ankle Tap **High Knee** **Knee Kick** **Reverse Lunge** **Side to Side**

Fig. 10. Adopted exercise motions in the experiment with their original videos and converted 3-D motions.

The platform supports the link to the Arduino board to collect the real IMU data and automatically fine-tune the model. Subsequently, the model is converted to the TensorFlowLite Micro model and can be loaded directly into an Arduino or Android device for applications.

IV. Evaluation

In this section, the evaluation have been conducted to validate the effectiveness of the proposed virtual IMU-based HAR system development platform. The first experiment was performed to evaluate the quality of the generated virtual IMU data. Another two main sub-experiments were involved in testing not only the effect of the proposed virtual IMU data augmentation method but also the case studies for different application scenarios with the proposed platform.

A. Dataset Collection

Since the platform is dedicated to supporting the virtual IMU data generation for the wearable HAR system development, there are few datasets for this purpose validation because the mainstream datasets only involve real IMU data. Thus, we established the IMU dataset for the proposed platform evaluation. In order to get close to the practical usage scenarios of the platform, we focused on a typical application, namely exercise tracking, which is commonly used in rehabilitation, daily sports, and gym. Exercise tracking normally requires a certain extent of customization since each user may have different motion habits and training targets. We selected five types of exercise motions. We downloaded the exercise video from YouTube and converted the exercise motion into a motion sequence file via an online tool (i.e., DeepMotion [45]). The involved motions are shown in Fig. 10. Each motion sequence was 25 s in length.

The 3-D avatar's motion is utilized to produce the virtual IMU dataset and train the CNN model. To validate the performance of the model, seven people (2 female and 5 female; average age 24.7) were recruited to generate the corresponding real IMU data. Five real IMU sensors (Xsens Dot [46]) were placed on the lower body locations of each individual, including the *right upper leg*, *right lower leg*, *left upper leg*, *left lower leg*, and *pelvis*. The sampling rate for

TABLE I
COMPARISON BETWEEN VIRTUAL IMU AND REAL IMU DATA REGARDING DTW DISTANCE

IMU Signal Type		Ankle Tap	High Knee	Knee Kick	Reverse Lunge	Side to Side
Virtual IMU vs. Real IMU	x	174.11 ± 92.08	154.62 ± 81.46	120.01 ± 101.18	76.48 ± 43.81	153.23 ± 41.36
	y	106.46 ± 50.60	81.78 ± 40.57	64.63 ± 28.05	60.67 ± 32.48	95.77 ± 46.99
	z	109.02 ± 45.14	110.72 ± 36.23	101.11 ± 83.94	63.42 ± 51.46	69.06 ± 35.46
Real IMU vs. Real IMU	x	185.12 ± 47.55	152.61 ± 50.25	214.17 ± 75.97	73.49 ± 11.39	96.92 ± 10.36
	y	103.08 ± 18.88	75.43 ± 16.83	65.56 ± 13.92	66.62 ± 23.58	95.21 ± 16.81
	z	104.79 ± 15.37	84.75 ± 8.29	144.87 ± 55.12	87.55 ± 30.27	64.47 ± 15.40

real IMU sensors was 20 Hz, which is the same as that of virtual IMU data. Each person has watched the motion exercise video to understand the body movement. Each motion was recorded in 90 s. Considering the practical usage, only one IMU sensor placed on the *right upper leg* was employed during the evaluation in this article.

B. Experiment 1: Evaluation of IMU Signals From 3-D Avatar Motion and Real Human Motion

Since virtual IMU data trains the machine learning model proposed by the platform to recognize the real IMU data eventually, we first evaluated the distribution of virtual IMU and real IMU data. To compare the two types of signals, we considered the dynamic time warping (DTW) distance as the metric, which is able to measure the similarity between two signals in the time domain [47], [48]. We used the augmented virtual IMU data from the proposed method to calculate the relative DTW distance with the real IMU data from each individual. For the baseline, we also tested the DTW distance between each individual's real IMU data as the baseline to gain insight into differences between the virtual IMU and real IMU datasets.

Table I gives the results of a comparison between the virtual and real IMU data. Before calculating metrics, we resampled the virtual and real IMU data to enable the time series are of the same length and obtain the results from each axis. From the results, the virtual IMU data did not present major differences in distribution compared with real IMU data, and different motions may lead to various results with different axes. Though the difference between the virtual IMU and real IMU is not significant compared with real IMU data, it is clear that the virtual IMU data have a more diverse signal distribution as the standard deviation of the virtual IMU data is bigger compared to that of the real IMU data. Since the machine learning model does not have prior knowledge of the real IMU data distribution during the pretraining period, expanding the distribution of the training data of the virtual IMU data is necessary. From Table I, we confirm that the similarity between virtual IMU and real IMU signals is acceptable and feasible for the pipeline of the proposed platform, which is also proved in the previous works [18], [19].

C. Experiment 2: Motion Modification-Based Virtual IMU Data Augmentation

The proposed virtual IMU data augmentation method reduces the requirement for the number of input avatar motion sequences, which can further contribute to the development of end-to-end user-oriented customized HAR systems. Therefore, we evaluated the effectiveness of the proposed virtual IMU data augmentation method on recognition model training. We tested the performance of the augmentation method compared with baseline methods, i.e., only employing the original virtual IMU for training. The experiment tested three, four, and five classified motion exercises, respectively. Three different machine learning models were also involved, including the RF with handcrafted features [49], the CNN model without the fine-tuning process, and the proposed model training approach in Section III-D. After the virtual IMU data augmentation, each motion's virtual data length was expanded from 25 s into 25 s $\times$ 4 factors $\times$ 8 filters = 800 s. Each individual's real data was split into four folds. The classifiers are trained using virtual IMU data and tested on real IMU data for RF and CNN models. The CNN with fine-tuning is performed using a few-shot real IMU data (1 fold/22.5 s) on fine-tuning and tested with the remaining real data (3 folds), which ensures no tested data leakage in the training phase.

The evaluation metrics of the classifier's accuracy, sensitivity, and specificity were calculated, which were common machine learning system evaluation metrics [27], [34]. The test repeated independent model training three times, and the results were averaged from the test [35].

Fig. 11 presented the results of data augmentation on virtual IMU data. For various numbers of motion categories, the hand-crafted features-based classifier (i.e., RF) did not show a good performance on recognizing the real IMU data with or without the data augmentation method. This is because the manually designed features usually presented limited representation ability on low-dimensional time series (e.g., IMU signal) and difficulty in realizing the transferring capability from the virtual domain to the real-world domain. For the CNN model, the proposed data augmentation is able to improve the performance when training with virtual IMU data and testing with real IMU data. Since the augmentation expands the richness of training data distribution, the CNN model could reach a higher accuracy when recognizing the real IMU data. Nevertheless, the fine-tuning strategy is capable of greatly improving the recognition performance of the virtual IMU data-trained CNN model. Compared with using the initial virtual IMU for training, the proposed data augmentation method could enhance the accuracy from 56.8% to 82.6% for three types of motions (c.f., the result of CNN with fine-tune in Fig. 11). With the recognized motion types increased, there is a corresponding decrease in the recognition performance of the model, e.g., 77.2% for four types of motion and 76.0% for

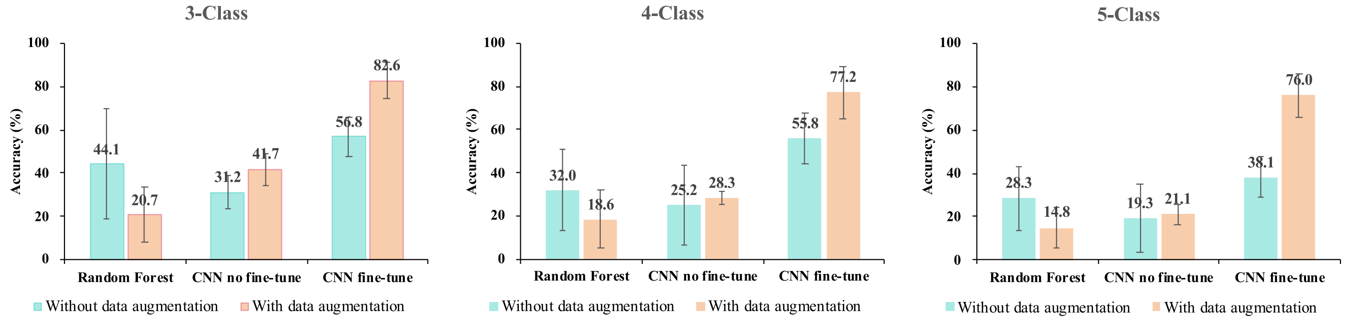


Fig. 11. Results of the effect of the proposed motion modification-based virtual IMU data augmentation method. The three recognized motions are *Ankle Tap*, *High Knee*, and *Knee Kick*. The four recognized motions are *Ankle Tap*, *High Knee*, *Knee Kick*, and *Reverse Lunge*.

five types. Therefore, it could prove the effectiveness of the proposed data augmentation in improving the performance of the virtual IMU data-trained CNN model.

D. Experiment 3: Case Study of Employing the Proposed Platform to Design a Customized HAR System

1) Personalized HAR System Development by Individuals: In the first case study, we assumed the scenario where the users developed the customized HAR for personal usage. The user could follow the virtual IMU-based HAR system's development of a pipeline from avatar motion sequence file import. Then, the platform will generate the dataset for the model training and collect a few shots of personal real IMU data to fine-tune the recognition model.

To validate this case, we evaluated using the individual's real data to fine-tune the pre-trained model and tested the model on the same user. Various amounts of individual's real IMU data used for fine-tuning were tested. A total of seven users' data were involved, and the result was calculated by averaging all the users' results. Also, we set the baseline method as the conventional HAR system building approach, which utilized the same size of real data to develop the machine learning system and test the remaining data from the same user.

Fig. 12 shows the result of the experiment. The effect of different amounts of real IMU fine-tuning data from individuals on the model performance was tested for different numbers of motion categories. From the figure, as the amount of real IMU data used increases, the recognition accuracy is raised. Compared with training the model with real IMU data in the traditional way, the proposed customized HAR system-building approach could outperform the common approach where real data trains the model. In this method, the user is asked to provide a portion of the data for fine-tuning. According to the experiment's results, the user only needs to provide a small amount (20–30 s) of real IMU data to realize the recognition of the corresponding real motions (over 80%) in three types of motion recognition. The greater the number of recognized motions, the more real data from the user will be needed for fine-tuning to achieve better accuracy.

Table II presents the sensitivity and specificity of the trained classifier. Overall, all the results from the proposed method outperform the baseline method. As the number of recognized activities increased, the various performance metrics

TABLE II
RESULTS OF ACCURACY, SENSITIVITY, AND SPECIFICITY IN PERSONALIZED HAR SYSTEM DEVELOPED BY INDIVIDUALS

Class Number	Method	Accuracy	Sensitivity	Specificity
3-Class	**Proposed**	80.3% ±13.9%	73.1% ±8.8%	84.7% ±5.6%
	Baseline	58.5% ±8.9%	50.7% ±16.2%	70.4% ±10.7%
4-Class	**Proposed**	73.5% ±14.4%	61.2% ±12.4%	82.9% ±7.2%
	Baseline	44.6% ±10.9%	46.9% ±9.3%	74.4% ±6.2%
5-Class	**Proposed**	66.9% ±10.0%	66.7% ±10.5%	89.4% ±4.8%
	Baseline	44.5% ±11.1%	44.5% ±12.6%	78.3% ±7.5%

decreased. Significantly, the sensitivity is slightly lower than the specificity in different numbers of activity recognition tests. The sensitivity provides a true positive rate metric and is also recognized as the recall. And the specificity is related to the performance of true negative prediction. Thus, it is notable that the machine learning model trained by virtual IMU data performs better on negative class prediction. This also revealed that there is still a certain extent of difference in the distribution of virtual and real IMU data distribution, while the inner difference of various activities' IMU data has a better performance in predicting the negative classes.

2) General HAR System Building: Case A presents the situation in which the individual used the personal data to fine-tune and develop the wearable HAR system for the individual self. In addition, we also tested using the partial real IMU dataset with all individuals to fine-tune the pre-trained model. This case simulated the scenario in the HAR, which was applied to a group of users and focused on the model's generalizability. Thus, the leave-one-subject-out evaluation was adopted, and the model's performance was tested under the different numbers of recognition activity categories as well.

Fig. 13 presented the general HAR system building result via the virtual IMU data training and real data for fine-tuning. Overall, using the virtual IMU data for pre-training and real IMU data for fine-tuning shows prior performance compared to using only the real IMU dataset for classifier building. For five types of motions, the proposed method could reach 90.6% accuracy, and the related baseline method is about

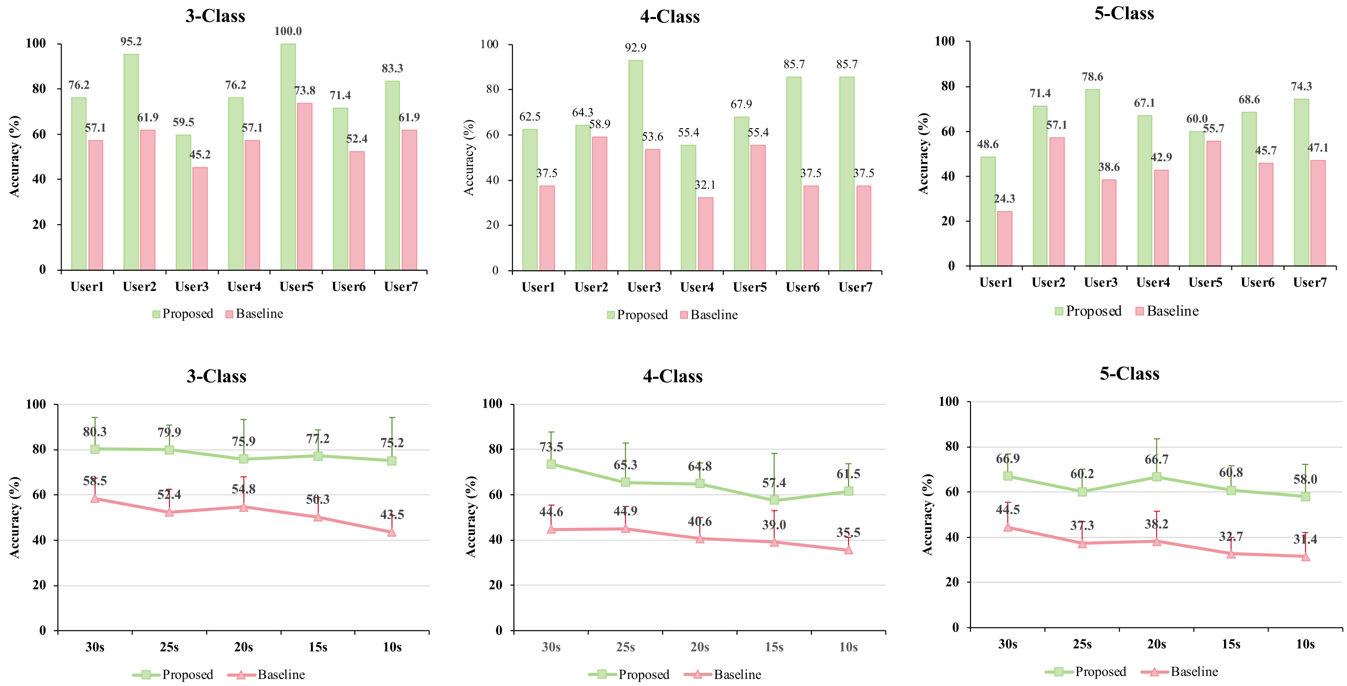


Fig. 12. Results of testing different amount of real IMU data on fine-tuning when training a customized HAR.

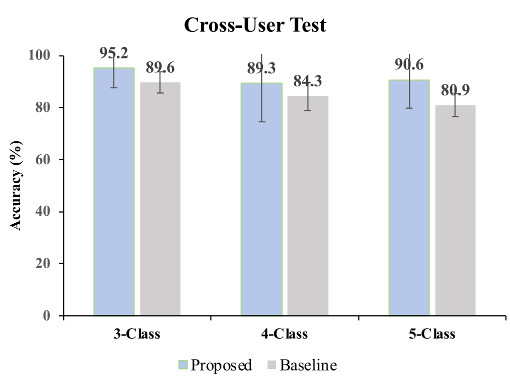


Fig. 13. Results of testing the generalizability of the model.

TABLE III
RESULTS OF ACCURACY, SENSITIVITY, AND SPECIFICITY IN GENERAL HAR SYSTEM BUILDING

Class Numebr	Method	Accuracy	Sensitivity	Specificity
3-Class	**Proposed**	95.2% ±7.6%	99.8% ±0.4%	99.9 % ±5.6%
	Baseline	89.6% ±4.1%	87.3% ±7.1%	93.2% ±4.3%
4-Class	**Proposed**	89.3% ±14.4%	93.4% ±9.7%	97.5% ±3.9%
	Baseline	84.3% ±10.9%	86.6% ±6.7%	94.9% ±2.8%
5-Class	**Proposed**	90.6% ±10.9%	92.6% ±9.3%	97.9% ±2.7%
	Baseline	80.9% ±4.4%	84.1% ±3.6%	95.7% ±1.1%

80.9%. For fewer recognized motions (e.g., three types), the proposed method is able to achieve 95.2%, while the baseline is 89.6%. In other words, the proposed method utilizes more training data than the traditional method of only utilizing real IMU data. Moreover, the additional IMU data involved was generated from the original 3-D motion sequence with a 25-s length, demonstrating the characteristic of low computational resources.

Table III shows the evaluation results. From the results, it can be seen that the classifier achieves better classification and is stronger than the baseline method, regardless of the task of recognizing several types of actions. All the methods used in the platform achieve good recognition ability with a higher accuracy than 90%. Both sensitivity and specificity maintain good performance. This also confirms that the use of virtual IMU data can effectively augment the training data distribution in the general wearable HAR system building. The trained classifiers have better generalization performance compared to using only real IMU data.

V. DISCUSSION

To facilitate a more convenient and efficient HAR system development paradigm, this article presents a platform for virtual IMU data-based HAR system building. The presented platform is open sourced in GitHub.[1] We introduced an end-to-end engineering pipeline to utilize the virtual IMU data for the wearable HAR system building, and it offers an alternative engineering approach to designing a prototype of HAR. The developer is capable of uploading the 3-D avatar motion sequence files, providing the personal few-shot real IMU data to obtain an employable machine learning model for practical deployment.

[1]GitHub Link: https://github.com/kannong/vIMU-HAR/tree/master

A. Virtual IMU Simulation Environment

In a physical environment, the accelerometer in a real IMU sensor utilizes MEMS technology to fabricate a tiny mass connected to a reference system by a spring, which is able to perceive the acceleration. In this article, we introduced the virtual IMU generation from the [36], and the whole simulation process did not follow the principle of a real IMU sensor's mechanism. The main simulation of the virtual IMU data is derived from a kinematic perspective based on the second displacement differential. This is because we are prone to gain the motion skeleton series from the 3-D avatar. Calculating the second differential of avatar joint displacement is much more intuitive and operable than simulating the mass-spring structure of a real accelerometer in an IMU sensor. It would decrease the requirement of the simulation environment in terms of physical principles, calculation resources, and so on. Thus, unlike other professional modeling or simulation software, e.g., MATLAB, Ansys, and others, our platform supports the normal virtual environment, such as Unity3D, to allow a more general interactive scenario, including the VR/AR applications. It emphasizes the iterative ability and low-threshold requirement of users.

B. Virtual IMU Signal Quality

Considering the difference between the virtual and real IMU data, the real IMU sensor usually generates a corresponding high-frequency noise during the measurement process due to the influence of the sensor structure and the transmission process. Notably, some high-frequency noise from the real IMU data and transmission loss is difficult to simulate with this approach. Therefore, considering the practical usage, the noise signal could be filtered, and this fine-grained simulation is not necessary for data-driven application scenarios such as those of the wearable HAR system. Moreover, since the virtual IMU data is derived from the 3-D motion in real-time, the hardware requirement to ensure a smooth and stable 3-D motion execution environment is significant, e.g., a relatively higher frame rate.

Additionally, the current works related to the virtual IMU follow the kinematic calculation approach to derive the virtual IMU data. Nevertheless, the data augmentation of virtual IMU data can also improve the quality of the virtual IMU dataset. For example, Xia and Sugiura [18] proposed a spring model structure in Unity3D to enhance the size of the virtual IMU dataset. Though the nonlinear principle of the spring model could increase the data distribution of virtual IMU, it still lacks an explanation, which makes it hard to control the related parameters. Thus, improving the quality of virtual IMU through the augmentation or modeling method, e.g., multidomain modeling, is still a promising solution to raise the fidelity of virtual IMU data.

On the other hand, the mathematical kinematics calculation approach to obtain the virtual IMU data may greatly depend on the avatar's motion input, which requires a reliable and stable 3-D motion as the premise as well as the IMU's wearing orientation. Since the virtual IMU data is used to train the HAR model, it is necessary to ensure a higher similarity between the training data and the ultimate recognized data distribution. So far, the platform has adopted transfer learning to bridge the gap between the virtual and real IMU data distribution. In addition, since there is no relevant data distribution knowledge for reading IMU during the pre-training process, it is significant to consider more virtual IMU data distribution. This is why there is a need to incorporate virtual IMU data augmentation methods in the platform. Considering the enrichment of virtual IMUs for data enrichment and increasing the degree of similarity between virtual IMU and real IMU data are key to enhancing the generalization capabilities of the model.

C. Deviation Between Virtual Avatar Motion and Real Motion

Our platform adopts the 3-D avatar's motion as the input reference to the real human motion from an RGB video conversion online tool [45]. However, there is always an inherent difference between the avatar's motion and real human motion. The conversion process follows an inference approach that employs a data-driven machine learning model to derive the 3-D skeleton information of the person in the video. Thus, the quality of the video would affect the completeness of avatar motion, such as the occlusion, lighting, background color, and so on [50]. However, the related evaluation of the conversion model from RGB video to 3-D avatar motion has been assessed in various ways to demonstrate its stability [51], [52].

However, the platform contributes by proposing an end-to-end pipeline from 3-D avatar motion to wearable HAR system building. Therefore, in this article, we did not evaluate how different the 3-D avatar motion is from the real motion. Because our goal is to utilize the 3-D avatar motion for generating the related virtual IMU data, and the motion's difference would result in the IMU signal's difference, we conducted the corresponding experiment to provide the result of signal-level evaluation to show the reliability of the proposed method. In addition, the approaches generating the 3-D avatar motions are diverse, including MoCap equipment, manual design, video conversion, and others. We deployed one of the common and accessible approaches, i.e., using the RGB video conversion. It is foreseeable that as the stability of the video conversion model improves, the quality of the resulting virtual IMU will inevitably become higher and higher.

D. Compared With Existing HAR System Development Platforms

Table IV presents the comparison of the proposed platform with previous toolkits/platforms for wearable HAR system development. Overall, all platforms basically concentrated on the machine learning route and supported the fundamental steps of system building, i.e., the signal collection and annotation, as well as the model training and testing. These platforms build interfaces to facilitate data collection via various communication protocols and support the visualization of data processing. In addition, the model training function helps the developer understand the model's performance and

TABLE IV
COMPARISON OF OUR PLATFORM WITH PREVIOUS TOOLKIT/PLATFORM WEARABLE HAR SYSTEM DEVELOPMENT

Work	Device	Develop Environment	Application	Training Dataset	Human Resources Cost	Sensor Number and Location Setting	Data Collection/ Annotation	Model Training/ Testing
[17]	IMU	Matlab	General HAR system development	Real	High		✓	✓
[38]	IMU	HTML	Support user experiment with COTS wearables	Real	High		✓	✓
[39]	EMG	Python	Interface prototyping and experiment	Real	High		✓	✓
[54]	IMU Camera	/	Gesture design	Real	High		✓	✓
[55]	IMU	JavaScript	General HAR system development	Real	High		✓	✓
[56]	IMU Camera	/	Gesture recognition system	Real	High		✓	✓
Our work	**IMU**	**Unity Game Engine**	**HAR prototyping and experiment**	**Virtual**	**Low**	✓	✓	✓

adjust the related configuration to realize an optimal recognition system. These two main steps remain in our platform as well. Basically, the evaluation result is similar to that of other experiments in HAR platform work. For example, Haladjian [17] recruited seven individuals contributing to the IMU dataset with eight classes of activity, and their platform built the wearable HAR system with 81.8% of accuracy.

From the comparison, our platform maintains the advantages in sensor number and location selection and has few real individuals participating in dataset collection to keep low human resources costs. First, the primary data generation process is performed by the 3-D avatar motion in the platform rather than the traditional real human. The training dataset is related to the virtual data, and thus, the requirement of participants to contribute to the training dataset is low. This would lead to a low-cost, high-efficiency solution, since there is no need to recruit the people and request them to repeat the recognized motion many times to generate the training dataset as the traditional way. Second, the platform gives more flexibility for the developer to indicate the sensor number and locations on the body to obtain the related IMU dataset. This is difficult to achieve on other HAR system platforms based on real IMU. This is because other platforms are based on a pre-collected dataset, which provides optional sensor locations and quantities. Once developers have new requirements, they need to collect datasets again. We converted the dataset collection from the real world into the virtual environment. Developers' new requirements can be met quickly, and the required sensor locations and quantities can be selected intuitively. Overall, with the proposed platform, the whole development process could be completed in the virtual engine. Only a few shots of real IMU data are required to complete the domain transfer and ensure a good recognition ability for the developed model. Therefore, the corresponding low-cost, high-efficiency, and flexible characteristics are the main merits of our wearable HAR system development platform.

VI. LIMITATION AND FUTURE WORK

A. Virtual IMU-Oriented Data Augmentation Method

The proposed platform introduces a motion modification-based virtual IMU data augmentation method. It aims to address the issue of larger resource requirements for 3-D motion sequences and lower the threshold of imported 3-D motion length. The evaluation demonstrated the performance of the proposed virtual IMU data augmentation approach. Using this method, the initial virtual IMU dataset size can be expanded, which will help improve the classifier's performance. However, there is still a lack of suitable public 3-D motion datasets with the corresponding real IMU data in the community. The most relevant dataset, such as AMASS, has many 3-D motions but still misses the real IMU data, which is difficult to employ to evaluate the platform's performance [53]. Thus, in this article, we only adopted a few motion exercises as the target recognized motions, and each motion's length is relatively short. Nevertheless, the involved motions have validated the developed platform's potential and effectiveness to a certain extent. In the future, building a suitable public 3-D motion with the corresponding real IMU data could benefit the investigation of data augmentation methods and classification algorithms.

B. Transfer Learning Approach for Virtual IMU-Based HAR System

Additionally, one of the most important challenges employing the classifier trained by virtual IMU data is facing the domain adaption problem, namely recognizing the real IMU data. In this article, we adopted few-shot real IMU data to fine-tune the dense layers of the CNN model to perform the knowledge transfer. Focusing on the cross-domain classifier establishment approach remains an issue of concern [20]. Since during the pre-training process, there is no prior knowledge in terms of real IMU data distribution, it is difficult to apply the manual information to compensate the gap between the virtual and real IMU data distribution. Thus, the platform asks users to provide certain real data, and adopt a more practical approach to conduct the transfer learning. The convolutional layer is frozen to ensure the feature extraction capability, and fewer real data are utilized to retrain the fully-connected layer to accomplish domain adaption. Other approaches, such as using mutual information, incremental learning, and knowledge distillation, although they can improve the generalization ability of the model, still lack the practicality characteristic. Therefore, continuing to explore more practical and high-performance transfer learning approaches to realize the transfer from virtual to real domains is the next step that still needs to be focused on.

So far, the experiment was conducted with limited recognized activity categories and participant numbers. This is due to the performance of the recognition model trained by virtual IMU data. As introduced, the difference between the virtual and real IMU data distribution results in a limited generalizability capability of the recognition model. However, since the platform has the advantage of decreasing the requirement of a real IMU dataset, it is suitable for application for a small population (e.g., a family), which only needs a little demand for model generalizability performance. The highly effective and flexible characteristics enable the development to be much easier. It is not only under virtual IMU data-trained models that such migration needs to be addressed. Other work, such as speech recognition [31] capacitive sensing [33], also meets challenges when migrating pre-trained models to individual users with personalized models. Investigating the method to realize a fast and high-performance transfer learning approach would also be the next focus of the research.

VII. Conclusion

This article introduced an open-source platform based on the 3-D avatar and virtual IMU data to build a customized wearable HAR system. Traditional HAR system development follows the typical sensor-based machine learning chain and maintains tedious and costly trial-and-error steps, especially the dataset collection. Therefore, we integrated the cross-modal virtual IMU data generation with the developed motion modification-based data augmentation and training method to be able to build the customized wearable HAR system. The advantage lies in the ability to flexibly select different wearing positions and numbers and generate relevant synthetic training datasets. The domain adaptation approach can also be utilized to further reduce the need for real data, improving the efficiency and cost of system development. Based on the experiment, the platform supports different application scenarios, and the user can establish and prototype the wearable HAR system faster, cheaply, and more efficiently.

References

[1] Z.-X. Yin and H.-M. Xu, “A wearable rehabilitation game controller using IMU sensor,” in *Proc. IEEE Int. Conf. Appl. Syst. Invention (ICASI)*, Apr. 2018, pp. 1060–1062.

[2] M. Straeten, P. Rajai, and M. J. Ahamed, “Method and implementation of micro inertial measurement unit (IMU) in sensing basketball dynamics,” *Sens. Actuators A, Phys.*, vol. 293, pp. 7–13, Jul. 2019.

[3] Y.-T. Hwang, Y.-Q. Tung, C.-S. Chen, and B.-S. Lin, “B-spline modeling of inertial measurements for evaluating stroke rehabilitation effectiveness,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 4008–4016, 2023.

[4] Y. Li, W. Chen, J. Wang, and X. Nie, “Precise indoor and outdoor altitude estimation based on smartphone,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

[5] C. Crema, A. Depari, A. Flammini, E. Sisinni, A. Vezzoli, and P. Bellagente, “Virtual respiratory rate sensors: An example of a smartphone-based integrated and multiparametric mHealth gateway,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2456–2463, Sep. 2017.

[6] G. Cosoli, L. Antognoli, L. Panni, and L. Scalise, “Indirect estimation of breathing rate using wearable devices,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–8, 2024.

[7] Y. Wu, J. Zhang, Y. Chen, J. Wang, W. Shi, and Q. Zhang, “Ubi-asthma: Towards ubiquitous asthma detection using the smartwatch,” *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11576–11587, Jul. 2023.

[8] F. Gu, M.-H. Chung, M. Chignell, S. Valaee, B. Zhou, and X. Liu, “A survey on deep learning for human activity recognition,” *ACM Comput. Surveys*, vol. 54, no. 8, pp. 1–34, 2021.

[9] T. Plötz, “If only we had more data!: Sensor-based human activity recognition in challenging scenarios,” in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops Other Affiliated Events (PerCom Workshops)*, Mar. 2023, pp. 565–570.

[10] W. Chen, S. Lin, E. Thompson, and J. Stankovic, “SenseCollect: We need efficient ways to collect on-body sensor-based human activity data!” *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 3, pp. 1–27, Sep. 2021.

[11] C. Tong, J. Ge, and N. D. Lane, “Zero-shot learning for IMU-based activity recognition using video embeddings,” *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 4, pp. 1–23, Dec. 2021.

[12] H. S. Ganesha, R. Gupta, S. H. Gupta, and S. Rajan, “Few-shot transfer learning for wearable IMU-based human activity recognition,” *Neural Comput. Appl.*, vol. 36, no. 18, pp. 10811–10823, Jun. 2024.

[13] X. Hu, Z. Wang, H. Weng, and X. Zhao, “Self-calibration of tri-axis rotational inertial navigation system based on virtual platform,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[14] X. Jin, S. Guo, A. Song, P. Shi, X. Li, and M. Kawanishi, “A novel robotic platform for endovascular surgery: Human–robot interaction studies,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–9, 2024.

[15] X. Allka, P. Ferrer-Cid, J. M. Barcelo-Ordinas, and J. Garcia-Vidal, “Temporal pattern-based denoising and calibration for low-cost sensors in IoT monitoring platforms,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

[16] J. Françoise, B. Caramiaux, and T. Sanchez, “Marcelle: Composing interactive machine learning workflows and interfaces,” in *Proc. 34th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2021, pp. 39–53.

[17] J. Haladjian, “The wearables development toolkit: An integrated development environment for activity recognition applications,” *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 4, pp. 1–26, Dec. 2019.

[18] C. Xia and Y. Sugiura, “Virtual IMU data augmentation by spring-joint model for motion exercises recognition without using real data,” in *Proc. ACM Int. Symp. Wearable Comput.*, Sep. 2022, pp. 79–83.

[19] H. Kwon et al., “IMUTube: Automatic extraction of virtual on-body accelerometry from video for human activity recognition,” *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–29, 2020.

[20] L. Pei et al., “MARS: Mixed virtual and real wearable sensors for human activity recognition with multidomain deep learning model,” *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9383–9396, Jun. 2021.

[21] P. S. Santhalingam, P. Pathak, H. Rangwala, and J. Kosecka, “Synthetic smartwatch IMU data generation from in-the-wild ASL videos,” *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 7, no. 2, pp. 1–34, Jun. 2023.

[22] J. Li et al., “SignRing: Continuous American sign language recognition using IMU rings and virtual IMU data,” *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 7, no. 3, pp. 1–29, Sep. 2023.

[23] Z. Leng, Y. Jain, H. Kwon, and T. Ploetz, “On the utility of virtual on-body acceleration data for fine-grained human activity recognition,” in *Proc. Int. Symp. Wearable Comput.*, Oct. 2023, pp. 55–59.

[24] J. Su, Z. Wen, T. Lin, and Y. Guan, "Learning disentangled behaviour patterns for wearable-based human activity recognition," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 6, no. 1, pp. 1–19, 2022.

[25] H. Zilelioglu, G. Khodabandelou, A. Chibani, and Y. Amirat, "Semisupervised generative adversarial networks with temporal convolutions for human activity recognition," *IEEE Sensors J.*, vol. 23, no. 11, pp. 12355–12369, Jun. 2023.

[26] T. T. Um et al., "Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks," in *Proc. 19th ACM Int. Conf. Multimodal Interact.*, Nov. 2017, pp. 216–220.

[27] S. Xia, L. Chu, L. Pei, Z. Zhang, W. Yu, and R. C. Qiu, "Learning disentangled representation for mixed-reality human activity recognition with a single IMU sensor," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–14, 2021.

[28] C. Hu, Y. Chen, X. Peng, H. Yu, C. Gao, and L. Hu, "A novel feature incremental learning method for sensor-based activity recognition," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 6, pp. 1038–1050, Jun. 2019.

[29] C. Hou, S. Gu, C. Xu, and Y. Qian, "Incremental learning for simultaneous augmentation of feature and class," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 12, pp. 14789–14806, Dec. 2023.

[30] T. Zhao, G. Cao, Y. Zhang, H. Zhang, and C. Xia, "Incremental learning of upper limb action pattern recognition based on mechanomyography," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 103959.

[31] Z. Su, S. Fang, and J. Rekimoto, "LipLearner: Customizable silent speech interactions on mobile devices," in *Proc. CHI Conf. Human Factors Comput. Syst.*, Apr. 2023, pp. 1–21.

[32] X. Xu et al., "Enabling hand gesture customization on wrist-worn devices," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2022, pp. 1–19.

[33] B. Steuerlein and S. Mayer, "Conductive fiducial tangibles for everyone: A data simulation-based toolkit using deep learning," *Proc. ACM Human-Computer Interact.*, vol. 6, no. 6, pp. 1–22, Sep. 2022.

[34] Y. Shi, Z. Tian, M. Wang, Y. Wu, B. Yang, and F. Fu, "Residual convolutional neural network-based stroke classification with electrical impedance tomography," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[35] S. Wang, T. Zhang, Y. Li, P. Li, H. Wu, and K. Li, "Continuous hand gestures detection and recognition in emergency human–robot interaction based on the inertial measurement unit," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–15, 2024.

[36] A. D. Young, M. J. Ling, and D. K. Arvind, "IMUSim: A simulation environment for inertial sensing algorithm design and evaluation," in *Proc. 10th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, Jun. 2011, pp. 199–210.

[37] L. Huang and C. Xia, "ModifyAug: Data augmentation for virtual IMU signal based on 3D motion modification used for real activity recognition," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, May 2024, pp. 1–7.

[38] O.-A. Schipor and R.-D. Vatavu, "GearWheels: A software tool to support user experiments on gesture input with wearable devices," *Int. J. Human–Computer Interact.*, vol. 39, no. 18, pp. 3527–3545, Nov. 2023.

[39] J. Karolus et al., "EMBody: A data-centric toolkit for EMG-based interface prototyping and experimentation," in *Proc. ACM Hum.-Comput. Interact.*, vol. 5, May 2021, pp. 1–29.

[40] H. Ding et al., "A platform for free-weight exercise monitoring with passive tags," *IEEE Trans. Mobile Comput.*, vol. 16, no. 12, pp. 3279–3293, Dec. 2017.

[41] C. Xia, A. Saito, and Y. Sugiura, "Using the virtual data-driven measurement to support the prototyping of hand gesture recognition interface with distance sensor," *Sens. Actuators A, Phys.*, vol. 338, May 2022, Art. no. 113463.

[42] T. Reilly, *Ergonomics in Sport and Physical Activity*. Champaign, IL, USA: Human Kinetics, 2009.

[43] H. Zhang et al., "MaeFE: Masked autoencoders family of electrocardiogram for self-supervised pretraining and transfer learning," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–15, 2023.

[44] C. B. Kumar, A. K. Mondal, M. Bhatia, B. K. Panigrahi, and T. K. Gandhi, "Self-supervised representation learning-based OSA detection method using single-channel ECG signals," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–15, 2023.

[45] *DeepMotion*. Accessed: Mar. 1, 2025. [Online]. Available: https://www.deepmotion.com/

[46] N. Schlage, A. Kitzig, G. Stockmanns, and E. Naroska, "Development of a mobile, cost-effective and easy to use inertial motion capture system for monitoring in rehabilitation applications," *Current Directions Biomed. Eng.*, vol. 7, no. 2, pp. 586–589, Oct. 2021.

[47] W. Choi, J. Cho, S. Lee, and Y. Jung, "Fast constrained dynamic time warping for similarity measure of time series data," *IEEE Access*, vol. 8, pp. 222841–222858, 2020.

[48] T. T. Tin, N. T. Hien, and V. T. Vinh, "Measuring similarity between vehicle speed records using dynamic time warping," in *Proc. 7th Int. Conf. Knowl. Syst. Eng. (KSE)*, Oct. 2015, pp. 168–173.

[49] C. Xia, A. Munakata, and Y. Sugiura, "Privacy-aware gait identification with ultralow-dimensional data using a distance sensor," *IEEE Sensors J.*, vol. 23, no. 9, pp. 10109–10117, May 2023.

[50] H. Kwon, B. Wang, G. D. Abowd, and T. Plötz, "Approaching the real-world: Supporting activity recognition training with virtual IMU data," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 3, pp. 1–32, Sep. 2021.

[51] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7753–7762.

[52] J. Wang et al., "Deep 3D human pose estimation: A review," *Comput. Vis. Image Understand.*, vol. 210, May 2021, Art. no. 103225.

[53] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black, "AMASS: Archive of motion capture as surface shapes," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Jun. 2019, pp. 5442–5451.

[54] D. Ashbrook and T. Starner, "MAGIC: A motion gesture design tool," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2010, pp. 2159–2168.

[55] D. Bannach, O. Amft, and P. Lukowicz, "Rapid prototyping of activity recognition applications," *IEEE Pervasive Comput.*, vol. 7, no. 2, pp. 22–31, Apr. 2008.

[56] K. Lyons, H. Brashear, T. Westeyn, J. S. Kim, and T. Starner, "GART: The gesture and activity recognition toolkit," in *Proc. 12th Int. Conf. Hum.-Comput. Interact.*, Beijing, China. Cham, Switzerland: Springer, Jul. 2007, pp. 718–727.