Check for updates

Latest updates: https://dl.acm.org/doi/10.1145/3790110

PDF Download
3790110.pdf
17 March 2026
Total Citations: 1
Total Downloads: 0

Published: 16 March 2026

Citation in BibTeX format

RESEARCH-ARTICLE

Vsens: Incorporating XR into the Process of Collecting Virtual IMU Data

FENGZHOU LIANG

TIAN MIN

CHENGSHUO XIA

YUTA SUGIURA

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 10, Issue 1 (March 2026)
https://doi.org/10.1145/3790110
EISSN: 2474-9567

Vsens: Incorporating XR into the Process of Collecting Virtual IMU Data

FENGZHOU LIANG*, Keio University, Japan
TIAN MIN*, Keio University, Japan
CHENGSHUO XIA, Xidian University, China
YUTA SUGIURA†, Keio University, Japan

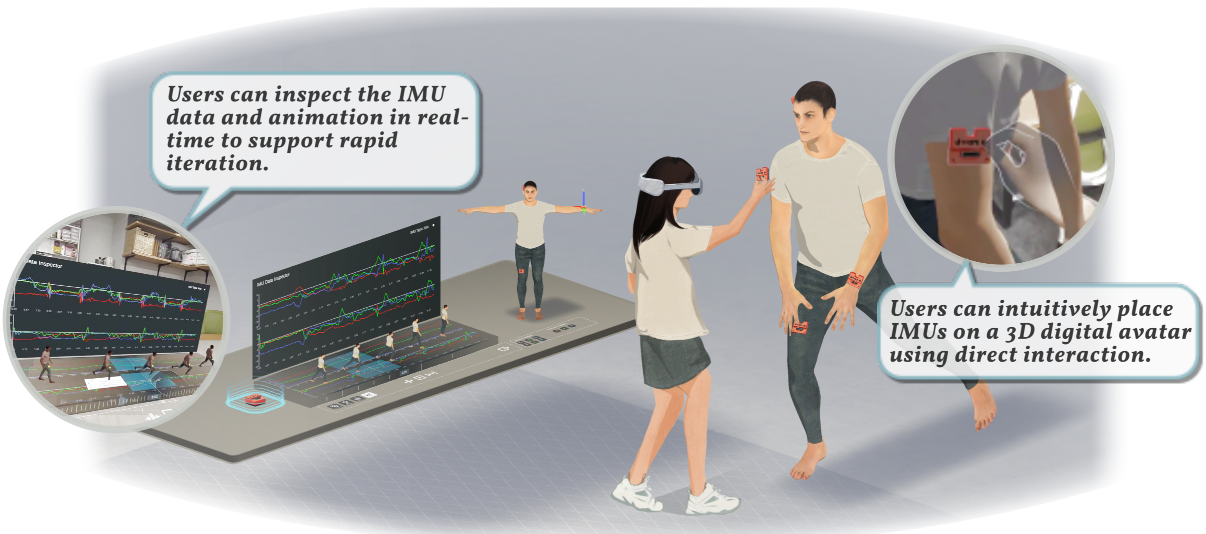


Fig. 1. We propose Vsens, an immersive virtual IMU data collecting interface that enables intuitive virtual IMU placement and adjustment.

Virtual Inertial Measurement Units (IMUs) offer a promising approach to generating synthetic motion data for training and evaluating human activity recognition (HAR) systems. However, existing virtual IMU workflows remain fragmented and technically demanding, requiring users to switch between 3D editors, scripting tools, and offline signal processing pipelines. These limitations hinder usability and iterative refinement for researchers and developers who design HAR systems. We present *Vsens*, an XR-based system that unifies virtual IMU configuration, visualization, and data synthesis within an immersive workspace. *Vsens* allows developers to directly manipulate sensor placements on digital avatars and observe synthesized

*Both authors contributed equally to the paper.
†Corresponding author.

Authors' Contact Information: Fengzhou Liang, Keio University, Yokohama, Japan, liangfz@keio.jp; Tian Min, Keio University, Yokohama, Japan, welkinmin@keio.jp; Chengshuo Xia, Xidian University, Guangzhou, China, xiachengshuo@xidian.edu.cn; Yuta Sugiura (corresponding author), Keio University, Yokohama, Japan, sugiura@keio.jp.

This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 2474-9567/2026/3-ART10
https://doi.org/10.1145/3790110

IMU signals in real time. To understand how developers with different expertise levels benefit from such an environment, we conducted a user study with 20 HAR developers covering a broad range of experience levels. Results show that *Vsens* improves configuration efficiency and fosters deeper spatial understanding of sensor behavior across the participants. Beyond usability, our findings reveal how embodied interaction and real-time feedback inform the design of practical virtual data collection systems that foster more effective and human-centered workflows.

CCS Concepts: • **Human-centered computing → Interaction techniques**; *Systems and tools for interaction design.*

Additional Key Words and Phrases: Extended Reality, Virtual Sensors, Toolkit, Data Collection

ACM Reference Format:
Fengzhou Liang, Tian Min, Chengshuo Xia, and Yuta Sugiura. 2026. Vsens: Incorporating XR into the Process of Collecting Virtual IMU Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 1, Article 10 (March 2026), 31 pages. https://doi.org/10.1145/3790110

1 INTRODUCTION

Recent advancements in embedded electronics and machine learning have accelerated the development of ubiquitous computing systems that integrate into our everyday environments. Among the sensing technologies, Inertial Measurement Units (IMUs) have become a core component in wearable devices due to their compact form factor, low power consumption, and ability to function independently of external infrastructure, making them particularly suitable for privacy-sensitive and always-available applications. IMU data has become indispensable to drive a wide range of applications, including human activity recognition (HAR), motion tracking [44], sports analytics [65], rehabilitation [11], and personalized healthcare monitoring [36]. As models grow more powerful and application scenarios become increasingly personalized or situated in edge cases, the demand for large-scale, diverse, and high-quality IMU datasets has increased, underscoring the importance of making IMU data collection more efficient and scalable to emerging contexts.

To address the scarcity of IMU datasets, recent research has increasingly turned to digital twin techniques. These approaches leverage virtual environments to simulate physical sensor behaviors, enabling cross-modal data collection pipelines of the virtual sensors. For example, by tracking the position of a point on the digital avatar over time, we can compute derived kinematic quantities such as velocity and acceleration, thereby synthesizing virtual IMU data [71]. This paradigm opens up the possibility of using large repositories of online videos or motion capture libraries in virtual IMU datasets, mitigating the limitations of real-world data collection. The virtual environment facilitates the generation of cross-modal sensor datasets (e.g., text-IMU [34], video-IMU [29], Doppler-IMU [4]) through shared spatial-temporal simulations. Another key benefit lies in the ability to systematically diversify and augment datasets by manipulating avatar morphologies [41], motion parameters [20], or sensor configurations [67]. The feasibility of applying virtual IMU data has been demonstrated in HAR tasks, including custom training scheme design [65], sign language recognition [35, 55], or accessibility for wheelchair users [36], suggesting its potential as a promising data generation method in the foreseeable future.

Despite the growing interest in using virtual IMUs for multi-modal data synthesis and dataset augmentation, a better practical process of generating virtual IMU data remains unexplored. Current workflows — typically involving a 3D editor (e.g. Unity, Blender) and external coding scripts (e.g. C#, Python) — lack interactive interfaces and real-time feedback mechanisms, resulting in time-consuming trial-and-error loops on virtual IMU position, orientation or axis calibration, which is similar to traditional IMU-based HAR development workflow in the real-world [48, 62]. Moreover, the lack of supporting tools imposes a significant barrier for developers who may wish to create personalized datasets [65], limiting the broader applicability and inclusiveness of virtual IMU-driven systems.

Extended Reality (XR), as an emerging interaction and display technology, has been shown to enhance prototyping [39, 70], visualization [40], and tutorials [23] in domains such as data analysis, education, and interactive modeling [5, 32, 64]. Its support for embodied spatial interaction makes it particularly well-suited for tasks that involve manipulating virtual content within 3D environments. In the context of virtual IMU development, XR enables users to physically reach into the space, place virtual IMUs on digital avatars, and adjust orientation through direct hand input. This interaction mode stands in contrast to traditional 3D editors, where configuration often involves indirect parameter tuning via scripting or mouse-based interfaces. Furthermore, XR facilitates rich visual feedback, allowing users to preview complicated spatial and temporal virtual IMU signals in real-time, compare them with reference signals, or iteratively refine their placements. These features make XR a promising foundation for interaction-centered and feedback-driven virtual IMU synthesis.

In this paper, we present *Vsens*, an open-sourced[1], immersive virtual data collection interface for HAR developers — practitioners involved in building, evaluating, or working with HAR systems, who possess a fundamental understanding of sensing modalities such as IMUs. This group typically includes researchers, engineers, and applied-domain professionals working with sensor data and modeling, who can expect to benefits from the virtual sensing technology. *Vsens* introduces a sensor-centric workflow that integrates virtual IMU configuration and signal preview into a unified pipeline. Developers can intuitively place IMUs on a 3D digital avatar using direct interaction, and manipulate the orientation through spatial widgets. Rather than focusing solely on signal generation, *Vsens* explores how immersive, embodied interfaces can transform the process of virtual data collection itself — turning it from a scripted pipeline into an interactive, feedback-driven activity. Our main contributions are as follows:

- We introduce *Vsens*, an XR-based system that unifies spatial interaction, visualization, and virtual IMU data synthesis into a coherent, feedback-driven workflow. This design lowers the cognitive and technical barriers in configuring virtual IMUs, making the process more exploratory and interpretable for HAR developers.
- We evaluate *Vsens* through a three-task user study with 20 participants (including early-stage researchers, faculty members, senior engineers, and medical doctors working in rehabilitation). The study covers a diverse yet domain-consistent group of HAR developers, providing insights into how the system supports varied workflows and usage contexts within this community.
- We outline key design implications for building practical virtual sensing tools and workflows. Our findings highlight how embodied interaction, real-time feedback, and integrated visualization can shape developers' understanding and iteration processes, pointing toward more human-centered approaches to virtual data collection and prototyping beyond *Vsens* itself.

2 RELATED WORK

To situate our work in the broader research landscape, we review previous work across three key areas: **virtual sensor data synthesis**, **XR interfaces**, and **human-centered tool for dataset generation**, where our research sits at the intersection of these domains.

2.1 Virtual Sensor Data Synthesis

Unlike domains such as computer vision or natural language processing, where benchmark datasets scale to millions of samples (e.g., ImageNet [10], One Billion Word Benchmark [8], the scale of labeled, on-body sensor datasets remains relatively small. Most public IMU datasets (e.g., PAMAP2 [53], MMFit [58]) contain only a few dozen participants and cover limited scenarios, due to the labor-intensive data collection process [9]. This limitation is especially pronounced in edge scenarios where real-world data is difficult, risky, or expensive to collect. For example, collecting data for fall detection, rehabilitation, or hazardous industrial tasks can pose challenges in safety, repeatability, or privacy. In this context, virtual sensor synthesis has emerged as a promising

[1]Available at: https://github.com/KEIO-LCLAB/Vsens

alternative without the cost of real-world instrumentation. Over the past decade, researchers have explored a wide range of modalities, including Doppler radar [1, 4], motion capture [51, 68], textual instructions [34], audio [37], and RGB video [29, 30], to support the creation of synthetic datasets for HAR systems.

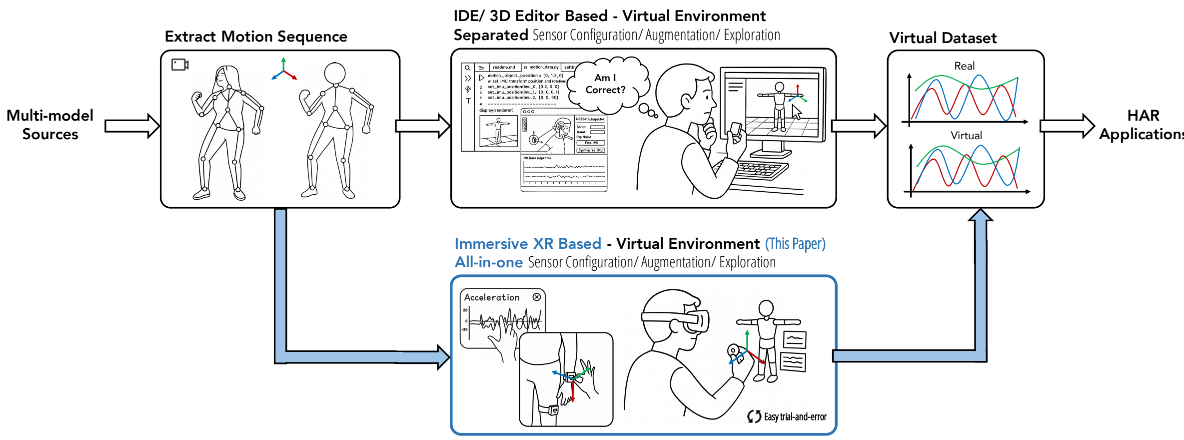


Fig. 2. Comparison between traditional and immersive workflows for virtual IMU data collection. The upper pipeline illustrates the conventional process based on IDEs or 3D editors, where motion data, sensor configuration, and dataset exploration are handled in separate tools and stages. The lower pipeline represents the proposed immersive XR environment (this paper), which unifies these steps into a single interactive workspace supporting rapid trial-and-error.

A typical virtual sensing pipeline begins by extracting human motion data from existing datasets, which is then mapped to a digital avatar within a virtual environment (Figure 2). Inside the 3D editors, virtual IMUs are instantiated as objects and manually placed onto the avatar mesh. These objects follow the avatar's movement and approximate the acceleration by computing the second derivative of position over time. Alternatively, the acceleration on a specific limb can be calculated through coding, where developers define the attachment joint, orientation offset, and sensor frame using scripts [71]. During the generation of virtual data, developers can diversify datasets through augmenting avatar parameters [41], motion speed [20], and style [21, 43]. However, virtual sensor data often suffers from alignment mismatches, animation glitches, or domain shift [42]. Developers commonly apply fine-tuning on a small set of real-world data to mitigate such issues.

While this pipeline has made significant progress in the collection of virtual sensor data, it often lacks dedicated tooling for spatial configuration, visual inspection, or human-in-the-loop refinement. Most existing pipelines treat virtual sensor setup as static or default, with limited elaboration for interactive exploration. Figure 2 summarizes this contrast: traditional IDE- or 3D-editor–based workflows fragment the process across multiple tools, separating motion import, sensor configuration, and dataset exploration into discrete stages. This fragmentation introduces friction for developers who must repeatedly switch contexts to inspect or adjust virtual IMU placements. In contrast, immersive approaches as proposed seek to unify these steps into a single environment, enabling easier iteration and embodied inspection. Section 4 provides a detailed description of how our system operationalizes these improvements through XR-based interaction and feedback mechanisms.

2.2 Assistive XR Toolkits

XR, as a bridging technology between physical and virtual environments, offers not only a highly immersive experience but also a set of unique affordances for spatial reasoning, embodied interaction, and multi-modal feedback. By enabling users to perceive, manipulate, and respond to virtual content from a first-person perspective, XR has become an increasingly valuable tool for interactive system design, rapid prototyping, and simulation-based workflows or tutorials.

In electronic prototyping, systems such as SchemaBoard [25] and CircuitStack [63] have addressed the challenges of physically wiring breadboards by providing integrated guidance and error prevention mechanisms. However, their hardware-centric nature makes them less adaptable to changes in electronic components [12]. XR-based prototyping toolkits offer greater adaptability across prototyping scenarios, as modifications to the software can support different components or sensing tasks without changes to the physical setup [33, 38]. While XR does not directly intervene in physical operations, it enhances developer awareness through richer information, supporting more flexible workflows.

A representative example that combines sensor-based prototyping with XR is SensorViz [26], which enables users to explore sensor specifications spatially in a 3D editor, visualize sensing fields in AR before fabrication, and inspect recorded sensor data on prototypes. It helps reduce trial-and-error in sensor selection and layout in the early stage of design. While SensorViz provides valuable insights into how XR can enhance sensor prototyping workflows, its design philosophy remains grounded in supporting physical development. Unlike visualization tools that passively display pre-collected signals, *Vsens* allows users to generate and manipulate data in real time, thereby transforming sensing from a post-hoc visualization problem into a controllable, interactive simulation process [66]. By decoupling prototyping from physical hardware, *Vsens* enables users to formulate and iterate on sensor placement strategies entirely within a virtual environment.

Another stream of XR-based toolkits highlights the power of immersive visualization for movement-centric tasks. For example, AvatAR [52] combines interactive 3D avatars with in-situ motion data visualizations to support spatial analysis. The use of embodied interfaces and mid-air interaction to navigate and explore detailed posture data demonstrates the effectiveness of XR in reducing cognitive load and enabling intuitive spatial reasoning. These findings align with our motivation to combine signal generation and motion feedback in *Vsens* and to support a sensor-centric, spatially grounded workflow for virtual IMU data collection.

2.3 Human-centered Tools for Dataset Creation

Recent years have seen increasing efforts to support non-expert users or researchers to better build machine learning datasets through human-centered interfaces, particularly in domains such as interactive machine teaching (IMT) [17, 54] and sensor data annotation [31, 76].

IMT systems [7, 75] allow users to demonstrate input-output mappings directly, often using real-time model feedback to guide labeling decisions. These systems lower the barrier to classification model construction but focus primarily on interaction-level behaviors rather than sensor configuration or data synthesis. In parallel, interactive annotation platforms aim to streamline the post-hoc exploration of time-series data. For instance, SegIt [73] supports temporal segmentation across multiple sensor channels, while M-MOVE-IT [24] provides an end-to-end framework for acquiring, synchronizing, annotating, and exporting multimodal datasets (e.g., IMU and video) for applications such as healthcare, sports, and animal monitoring. The strength of M-MOVE-IT lies in how it simplifies synchronization between sensors and videos, generates structured annotation tasks, and supports human-in-the-loop development. Its web-based interface provides users with tools to manage sensors, align video and IMU timelines, and annotate both modalities in tandem—empowering researchers to focus on meaningful data segments rather than tedious setups. These workflow enhancements highlight an important insight: even for real sensors, intuitive and well-integrated authoring interfaces are critical to scalable, high-quality dataset creation.

This insight applies equally to the virtual domain. Virtual IMU systems are often designed around automation or signal fidelity, with little attention to how users configure and inspect virtual sensors during dataset synthesis. *Vsens* fills a complementary role as a dataset creation tool positioned in the field of virtual sensing technologies.

3 CHALLENGES IN USING VIRTUAL IMU SENSORS

In this section, we conceptually bridge insights from real-world IMU challenges to their virtual counterparts, to articulate what remains unsolved in current virtual sensing workflows and why these gaps matter to HAR developers.

3.1 Bridging Real and Virtual IMU Limitations

Chen et al. [9] identified three major challenges in the process of collecting on-body sensor data for IMU-based HAR systems: **(a) Tedious and Time-consuming Process:** The setup is often tedious, laborious, and time-consuming, particularly for non-researchers or field deployments, due to device configuration, body placement, and calibration overhead; **(b) Lack of Interpretability:** The collected IMU signals are difficult to interpret or visualize, making it challenging to further annotate or validate activities based on raw IMU data; **(c) Absence of Feedback During Capture:** Without real-time feedback mechanisms, users cannot perceive how their motion is being captured. As a result, data is often polluted by unconscious behavior, deformations, or failed capture attempts. These challenges make large-scale IMU data collection highly dependent on expertise and prone to reliability issues. They also underscore the importance of interactive inspection and iterative refinement for any sensing workflow, whether physical or virtual.

In response, virtual IMU synthesis has emerged as an alternative approach, offering the ability to generate inertial data from motion sequences without the overhead of physical instrumentation. Virtual IMU largely alleviates the tedious burden of physical IMU attachment and calibration **((a) Tedious and Time-consuming Process)**, yet new bottlenecks appear in spatial configuration and inspection. In typical virtual IMU workflows environment (e.g. Unity-, Blender-, or script-based pipelines), the process is dominated by low-level operations — users must type coordinates, adjust transform values, or choose from a fixed set of pre-defined attachment points [9, 28]. The interface layer is often treated as an implementation detail rather than a design consideration. As a result, many virtual IMU tools overlook how developers actually interact with and reason about virtual IMU workflows. Even in recent publications [20, 61, 67, 71], the steps of positioning, orienting, or verifying virtual IMUs are either handled through code snippets or omitted altogether, revealing a gap between the algorithmic focus of prior work and the interactive demands of real-world HAR development.

While tedious setup has been recognized as a burden in real-world sensing **(a)**, its underlying difficulty often lies in achieving consistent placement and orientation across recording sessions, as pointed out by Banos et al. [3]. During HAR system development, small deviations in sensor position or orientation can lead to substantial differences in acceleration and angular velocity patterns, thereby altering model performance and generalizability [6, 18]. From a research standpoint, such sensitivity makes sensor placement a key controllable variable—precise and repeatable placement is essential for reproducibility, benchmarking, and fair model evaluation.

These insights extend directly to the virtual domain: the precision and consistency are critical for ensuring the validity of synthesized signals. A typical example is that, when evaluating against real-world datasets, such as PAMAP2 [53] or MM-Fit [58], inconsistencies between the virtual IMU's frame and the actual wearable setup can introduce subtle but consequential deviations that are difficult to detect and correct. In many widely used datasets, reference IMU signals (real IMU signals) are available, but the corresponding IMU placement is often described only coarsely or through minimal metadata, leaving orientation offsets or attachment styles unspecified. Admittedly, in real-world HAR deployment, sensor positions inevitably shift due to user motion or attachment differences. However, for system development and research, controlling placement as an experimental variable remains crucial—it allows developers to isolate model behavior, reproduce dataset conditions, and establish consistent baselines for cross-domain evaluation.

Moreover, fine-grained control over sensor configuration is increasingly desirable in modern HAR research. As wearable sensing moves beyond conventional attachment points such as wrists or thighs, researchers have begun to explore alternative placements (e.g., rings [49], necklaces [72]) to capture richer or more context-specific motion signals [74]. These emerging sensing form factors expand the design space of IMU-based systems, but they also make the configuration process more intricate: each placement may require distinct orientation alignment, signal calibration, and attachment modeling. Virtual sensing workflows, therefore, should not only replicate fixed sensor locations from existing datasets but also support flexible, user-defined configurations that reflect real-world diversity.

Prior work on real-world IMU data collection [9] highlights that limited signal interpretability makes downstream tasks such as annotation difficult and error-prone **((b) Lack of Interpretability)**. In virtual IMU workflows, although annotation may not be required in the same way, thanks to a highly integrated code environment and information available in the virtual scene, developers still face challenges in understanding how sensor placement and motion configuration translate into signal characteristics. Existing virtual IMU tools provide little systematic support for directly linking synthesized signals to observable motion events, leaving users to rely on low-level signal plots without spatial or temporal grounding [20, 61, 71]. Without synchronized visualization between animation and IMU signals, identifying configuration errors or validating intended activities remains tedious. In principle, virtual environments are well-suited to mitigate this challenge. Because both motion trajectories and sensor signals are generated within the same simulation space, they can be temporally aligned, spatially inspected, and replayed without the noise and uncertainty of real-world recordings. Developers can, in theory, observe how a particular limb movement produces corresponding accelerometer or gyroscope traces, or how a sensor's orientation affects the resulting signal.

Finally, real-time feedback during virtual IMU configuration remains limited **((c) Absence of Feedback During Capture)**. Most current pipelines operate offline: users must export motion sequences, synthesize signals, and inspect results across disjointed tools. This fragmented process delays feedback, making it difficult to detect misconfigurations early or refine IMU parameters iteratively. Yet surprisingly, real-time feedback is one of the aspects that virtual environments are most capable of providing. Because motion, signal, and configuration all originate from the same simulation space, a unified interface could in principle, allow immediate inspection of how each placement or parameter adjustment affects the synthesized IMU signal.

In summary, the transition from real-world to virtual IMU data collections does not eliminate the core challenges of placement, interpretability, and feedback. While virtual sensing removes the physical constraints of on-body instrumentation and is able to generate massive data easily (as described in Section 2.1), it inherits the same reproducibility and comprehension issues that hinder efficient HAR development. Despite its potential, the interface layer for virtual IMU collection remains largely unexplored: existing pipelines provide functionality for signal synthesis, but little design attention has been paid to how users actually configure, inspect, and iterate on sensors within immersive environments. To guide the design of such systems, we next summarize three key challenges distilled from the above analysis.

3.2 Key Challenges Informing System Design

Based on these observations, we identify three primary challenges that hinder the practical use of virtual IMU workflows in HAR development, and together they motivate the design goals presented in the next section (Section 4):

Challenge 1: Complex Interaction. Existing virtual IMU configuration typically involves abstract, code-driven interfaces or indirect manipulation in 3D editors. Developers must adjust transform gizmos or numerical parameters to manually align IMUs in 3D IDE, requiring mental translation of spatial relationships without embodied feedback. This limitation affects all HAR developers: junior developers struggle to build intuition about sensor behavior, while experts face increased time and effort to ensure reproducibility. For HAR development, this lack of spatial inspection interrupts the iterative process of calibration and validation, resulting in longer turnaround and reduced confidence in dataset integrity.

Challenge 2: Lack of Real-time Inspection and Feedback. Although virtual environments theoretically provide complete control over motion and signal generation, current virtual IMU workflows still offer little support for interpreting synthesized data in relation to observable movement. Most tools separate motion playback from signal visualization: avatars are animated in one interface, while acceleration and angular-velocity traces are inspected elsewhere as static plots. Developers must mentally map waveform patterns to body movements, whether developing new datasets, validating configurations, or debugging models. For HAR developers, the absence of synchronized visualization diminishes one of the key advantages of virtual sensing: the ability to directly observe cause-and-effect between motion and signal. Providing integrated, time-linked visualization would transform signal inspection from a detached analytic task into an interactive process.

Challenge 3: Reliable Virtual Data. Unlike the previous two challenges that originate from real-world sensing practices, this one arises from the intrinsic characteristics of synthetic data. The reliability of virtual IMU signals is a fundamental concern in practical HAR development. Because synthesized data are derived from motion models and simulation engines, their fidelity depends heavily on animation quality, body modeling, and numerical differentiation methods. Numerous studies have sought to enhance such realism through advanced physical simulation or domain adaptation [18, 20, 34, 42]; We do not attempt to solve data reliability itself, which has been widely addressed in other research. Rather, we acknowledge it as a design requirement of a virtual sensing system: remain open and compatible with diverse data sources and augmentation methods. Providing standardized interfaces for importing, augmenting, and validating virtual IMU data allows developers to plug in future improvements without disrupting existing pipelines.

Together, these challenges outline the core usability gaps in current virtual IMU workflows and inform the design rationale of our system. In the next section, we present *Vsens*, which was developed to address these challenges through an integrated, feedback-driven approach.

4 VSENS SYSTEM

Vsens is an XR-based system designed to collect virtual IMU data, with a particular focus on enabling more usable and iterative workflows for HAR developers. It is important to note that the XR embodiment, integrated workflow, and the following design rationales are not intended as isolated solutions to the individual challenges identified in Section 3, but rather as interdependent components of a unified system that collectively explore what a more effective virtual sensing interface can be.

R1. Integrating Motion and Signal. *Vsens* provides a tightly coupled visual representation that overlays synthesized IMU signals with the corresponding motion of the avatar. This integration supports more intuitive problem-finding, allowing users to trace sharp signal transitions or anomalies back to specific motion events.

R2. Dynamic Virtual Data Synthesis. To support rapid iteration, *Vsens* dynamically updates sensor outputs in real-time as users interact with the system. When a sensor is moved, reoriented, or newly placed on the avatar, the corresponding signal is immediately re-synthesized and visualized. This tight feedback loop enables continuous adjustment within a single immersive workspace, reducing context-switching and facilitating exploratory design without the need to recompute data through external tools.

R3. Compatibility and Data Augmentation. *Vsens* is designed to accommodate a wide range of motion sources, including SMPL [41], SMPL-X [50], and FBX formats derived from motion capture pipelines or vision-based pose estimation (e.g., WHAM [56], BEV-based reconstructions [59]). By supporting SMPL-compatible avatars, *Vsens* also enables body-shape–conditioned augmentation, allowing users to visualize and validate sensor behavior across diverse anthropometric characteristics. This extensibility supports future integration with data augmentation pipelines, making *Vsens* applicable beyond basic configuration and into dataset scaling workflows.

4.1 Virtual IMU in Vsens

Vsens synthesizes six-axis virtual IMU data by computing linear acceleration and orientation on a digital avatar. The acceleration is derived as the second derivative of the global position of each virtual IMU over time [71], while the orientation is represented as both Euler angles and quaternions, calculated based on the global rotation of the virtual IMU. Users can specify the sampling rate and choose the types of signals to record. To ensure smooth motion representation of virtual IMU, the system applies linear interpolation to position offsets and spherical linear interpolation (SLERP) [57] to adjacent rotation frames. To mitigate artifacts caused by animation jitter or sudden motion discontinuities, users can optionally enable a smoothing window for post-processing. All synthesis parameters, including sampling frequency, output data type, and smoothing configurations, can be customized per sensor to support diverse prototyping and experimentation scenarios.

4.2 XR User Interface

As shown in Figure 3, the *Vsens* interface consists of six major components, labeled (1) to (6) in the figure. When the system launches, components (1–5) are automatically arranged on nearby planar surfaces, while a life-sized *Main Actor* (6), approximately 180 cm tall, is placed upright on the floor. Each interface component can be individually detached and relocated, allowing users to customize the workspace layout as needed.

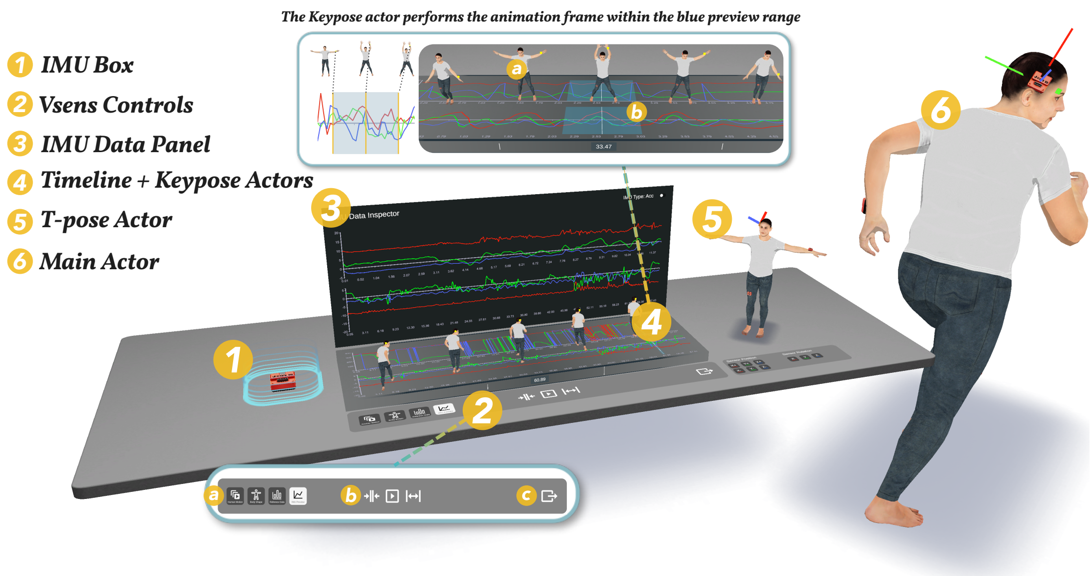


Fig. 3. *Vsens* interface includes six main components: IMU Box, *Vsens* Controls, IMU Data Panel, Timeline + Keypose Actors, T-pose Actor, and Main Actor.

Users can grab a new virtual IMU from the *IMU Box (1)* at any time. To prevent accidental activation, they must hold the IMU for 0.5 seconds; the IMU will then vibrate gently to indicate that it is active and can be freely moved or rotated using pinch gestures. Releasing the gesture drops the IMU, fixing it in place. When brought close to the *Main Actor (6)*, the IMU automatically snaps to the nearest mesh surface and is bound to the closest joint in the skeleton hierarchy.

To assist with inspection and multi-sensor placement, *Vsens* includes a static *T-pose Actor (5)* placed on a tabletop. This actor serves three purposes: (a) It allows users to inspect virtual IMU placements when the Main Actor is moving or partially occluded. (b) users can select which IMU signal is displayed on the *IMU Data Panel* (3) by pinching the IMU on the T-pose Actor; (c) It also provides a stable configuration point where users can adjust IMU positions and orientations more easily on a smaller scale.

Once motion data is loaded, the *Central Platform (4)* displays five *Keypose Actors (4.a)* sampled from a user-defined temporal segment in the animation. A blue overlay on the timeline determines the active range from which keyposes are selected. Each keypose is sampled evenly across the time span, providing a representative overview of the movement. Below the actors, a panel (4.b) displays rotation angles and angular velocities, offering temporal landmarks for motion understanding.

While the *Central Platform* (4) focuses on linking motion with signal interpretation, the *IMU Data Panel (3)* highlights the difference between synthesized virtual signals and reference IMU data. Users can toggle between acceleration and orientation visualization via a button in the top-right corner. Any changes to sensor placement, orientation, avatar morphology, or animation will be immediately reflected in the display. When switching between multiple virtual IMUs, the system automatically displays the matched reference signal, if available.

At the bottom of the workspace, the *Vsens Controls (2)* toolbar provides three button groups. The left group (2.a) allows users to switch the center platform between animation selection, body shape adjustment, and reference data visualization. The center group (2,b) provides playback controls and lets users set the time range for *Keypose Actors*. The right group (2.c) allows users to export current virtual IMU data to CSV, along with a scene configuration template. These templates can be shared across devices for batch virtual IMU data collection with predefined configurations (e.g., same IMU placement with different motion animations).

4.3 Use Case Examples

In this section, we present four major steps to collecting virtual IMU data using Vsens, representing a basic *Vsens* workflow.

Actor Setup. In the initial stage of Vsens, users begin with the actor setup, which serves as the basis for IMU synthesis results. *Vsens* supports three animation file formats: SMPL [41], SMPL-X [50], and FBX. These files can be imported into the *Vsens* system by placing them in the project library. Figure 4.1, users can select a desired motion animation by clicking the settings interfaces of the *Vsens* controls. After selecting an animation, users have the option to adjust the body shape of the virtual actor by dragging sliders. These sliders control various anthropometric parameters obtained from the PCA result of SMPLX [22], enabling users to tailor the actor's physique to closely match the real-world individuals. In addition, *Vsens* provides settings to influence the virtual IMU data and enhance data diversity. As shown in Figure 4.1.c, users can adjust the simulation process by modifying parameters such as the sampling rate and smoothing window size. They can also modify the motion speed and motion amplitude to enhance the playback effects of the animation.

Virtual IMU Placement. After setting up the actor, users proceed to place virtual IMUs intuitively, as shown in Figure 4.2. Users can pinch an IMU sensor from the IMU Box and move it toward the Main Actor. For more precise adjustments, *Vsens* provides a static T-pose Actor, and UI controls that allow users to accurately rotate the IMU around specified axes and translate it along axis directions in 1 cm increments, enabling fine-grained sensor placement. To enhance usability, *Vsens* also allows users to adjust the position and rotation of the Main Actor using the controller's thumbstick. This functionality provides a flexible and convenient way to reposition and rotate the actor, facilitating optimal sensor placement and improved visibility from multiple angles.

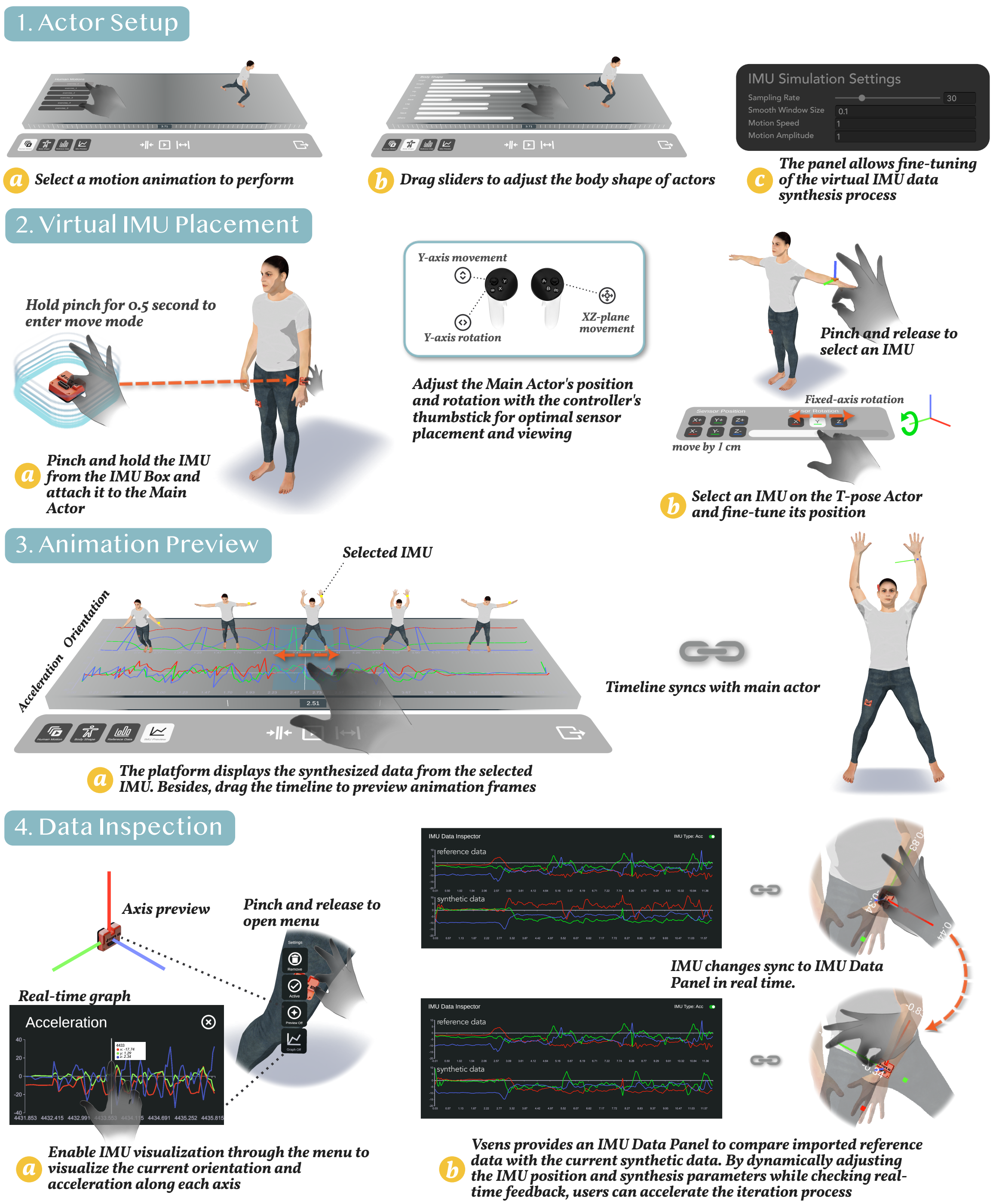


Fig. 4. Four main steps during the process of using Vsens: (1) Actor Setup; (2) Virtual IMU Placement; (3) Animation Preview; (4) Data Inspection.

Animation Preview. *Vsens* provides users with an intuitive animation preview by combining timeline, IMU data, and actors' animation. As shown in Figure 4.3, users can control the animation frame currently playing by dragging the indicator on the timeline, with changes synchronized to both the Main Actor and the Keypose Actors on the platform. The blue preview range and the five Keypose Actors help users better understand the motion details of the animation within a specified time window, as well as the corresponding IMU data.

Data Inspection. A common scenario when synthesizing IMU data is that users often know the signals from real IMUs but find it challenging to identify precise IMU placements to reproduce these signals accurately. *Vsens* facilitates this process through intuitive data inspection tools (Figure 4.4). Users can enable IMU visualization from the menu to intuitively observe the IMU's real-time orientation and acceleration along each axis, which aids users in building a clear understanding of how the placement influences IMU signal. Users can place reference IMU data in the project library and select it by clicking the settings interfaces of the *Vsens* controls. Enabling the IMU data panel to compare the imported reference IMU signal directly with the current synthetic IMU signal. Real-time feedback helps users swiftly iterate and refine IMU placements and synthesis parameters, enabling efficient alignment with target signals.

5 USER STUDY

We conducted a within-subject user study to evaluate the usability and effectiveness of *Vsens* for HAR developers, using three representative configuration and analysis tasks followed by semi-structured interviews. The study aimed to examine how our system supports sensor placement, signal inspection, and iterative refinement compared with existing virtual IMU workflows.

5.1 Hypotheses

We formulate the following hypotheses for our user study, which reflect challenges faced by HAR developers during virtual sensor configuration and validation (described in Section 3). These hypotheses guide the evaluation metrics and significance testing conducted in the following sections:

- **H1:** Accurate and consistent sensor placement is critical for reproducibility. In practice, developers must frequently reposition virtual IMUs to replicate experimental setups or explore new configurations across datasets. For example, developers may examine how moving a sensor from one place to another affects signal shape. We expect participants will be able to place virtual IMU sensors more efficiently when using *Vsens* compared to conventional virtual IMU workflows. **We expect reductions in task completion time and iteration count.** While spatial precision (transform error) is important, **we primarily hypothesize that *Vsens* will achieve comparable precision relative to conventional tools.**
- **H2:** Aligning virtual IMU orientation with reference signals is a routine yet time-consuming step in validating synthesized data. Although algorithmic methods can estimate orientation by minimizing signal discrepancies [27], they often require multiple iterations and lack transparency, making it difficult for developers to interpret or intervene during the process. Our study uses this task to examine the value of real-time feedback in facilitating virtual IMU configuration. **We expect participants will be able to more accurately and efficiently find the orientation of virtual IMUs through matching reference signals when using *Vsens*** compared to conventional virtual IMU workflows. We hypothesize that real-time signal feedback in *Vsens* will lead to lower signal deviation and faster convergence compared to offline adjustment or algorithmic optimization methods.
- **H3:** Virtual IMU data generated through *Vsens* can provide accuracy and representational quality comparable to real-world IMU datasets in typical classification tasks. **We expect that developers with diverse backgrounds will be able to utilize *Vsens* to produce data of sufficient fidelity for model training**

and validation. In other words, the system not only improves workflow usability but also yields virtual sensor data with practical reference value for HAR research and development.

5.2 Participants and Procedure

We recruited 20 participants (15 male, 5 female), aged between 22 and 41 years-old (*median* = 28), who represented a broad range of backgrounds relevant to HAR system development. All participants were practitioners involved in the design, prototyping, or evaluation of HAR-related systems. The sample included individuals from ***academia*** (11), ***industry*** (4), and ***clinical*** (5) domains, providing a view of both research and applied contexts. Their occupations ranged from Ph.D. candidates and master's students to R&D engineers and medical doctors specializing in biomechanics. As shown in Figure 5, participants' experience in programming, HCI, and HAR spanned from less than one year to over decades, capturing both junior developers beginning to explore HAR workflows and experts with extensive professional experience. Among the participants, PA-1, PA-4, PA-5,

ID	Age	Gender	Affiliation Type	Occupation	Discipline	Experience in (years)		
						Programming	HCI	HAR
PA-1	38	M	Academia	Faculty (Prof.)	HCI	15	15	15
PA-2	30	M	Academia	Ph.D.	System Design	10	2	2
PA-3	28	M	Academia	Ph.D.	System Design	10	<1	<1
PA-4	28	F	Academia	Ph.D.	HCI	15	>10	>10
PA-5	28	M	Academia	Ph.D. Candidate	HCI	10	5	5
PA-6	27	F	Academia	Ph.D. Candidate	Machine Learning	6	<1	<1
PA-7	27	M	Academia	Ph.D. Candidate	HCI	7	4	4
PA-8	26	M	Academia	Ph.D. Candidate	HCI	9	3	3
PA-9	25	F	Academia	Ph.D. Candidate	XR	4	3	<1
PA-10	23	F	Academia	M.S. Student	Data Science	3	<1	<1
PA-11	22	M	Academia	M.S. Student	HCI	3	<1	1
PI-1	31	M	Industry	R&D Engineer	Health	13	>10	>10
PI-2	29	M	Industry	R&D Engineer	HCI	15	6	3
PI-3	27	M	Industry	R&D Engineer	HCI	10	6	4
PI-4	26	M	Industry	R&D Engineer	Sensing	10	0	1
PC-1	41	M	Clinical	Medical Doctor/ Prof.	Biomechanics	0	0	>20
PC-2	39	M	Clinical	Medical Doctor/ Prof.	Biomechanics	18	5	>10
PC-3	33	M	Clinical	Medical Doctor	Biomechanics	4	0	>5
PC-4	32	F	Clinical	Medical Doctor	Rehabilitation	2	0	>5
PC-5	31	M	Clinical	Medical Doctor	Rehabilitation	0	0	3
Median						**9.5**	**2.5**	**3.5**

Fig. 5. Demographic information of the participants, denoted with *PA* (Academia), *PI* (Industry), and *PC* (Clinical) according to their affiliation types.

and PI-1 have experience working with virtual sensing technology. This diversity enabled us to analyze how developers at different levels of expertise perceive and benefit from immersive virtual IMU workflow.

Each participant received compensation of 30 USD for their time. The study protocol was reviewed and approved by the institutional ethics committee. Prior to participation, we confirmed that none of the participants had a history of cyber or motion sickness, and no such symptoms were reported during the study. For participants who wore glasses, we prepared a compatible lens adapter for the Meta Quest 3 headset and assisted with interpupillary distance adjustment to ensure visual comfort and safety throughout the experiment.

The study procedure began with a verbal briefing that introduced the goals and overall structure of the experiment. Participants then filled out a consent form and a demographic questionnaire. To ensure adequate background understanding, we provided printed guidelines that introduced core concepts of IMU-based sensing and virtual sensor technology. The study was conducted in a 3 m × 3 m open space, allowing for free movement during XR interaction. Each task lasted approximately 20 minutes, with a 10-minute break between tasks. The full study session took around 110 minutes. A researcher remained present throughout the session to assist with setup, ensure safety, and answer participant questions, with another researcher assisting with the device setup, data collection, and annotation. The order of study conditions was counterbalanced across participants using a Latin Square design. Prior to the formal study, we conducted an internal pilot session to refine task instructions and verify hardware stability.

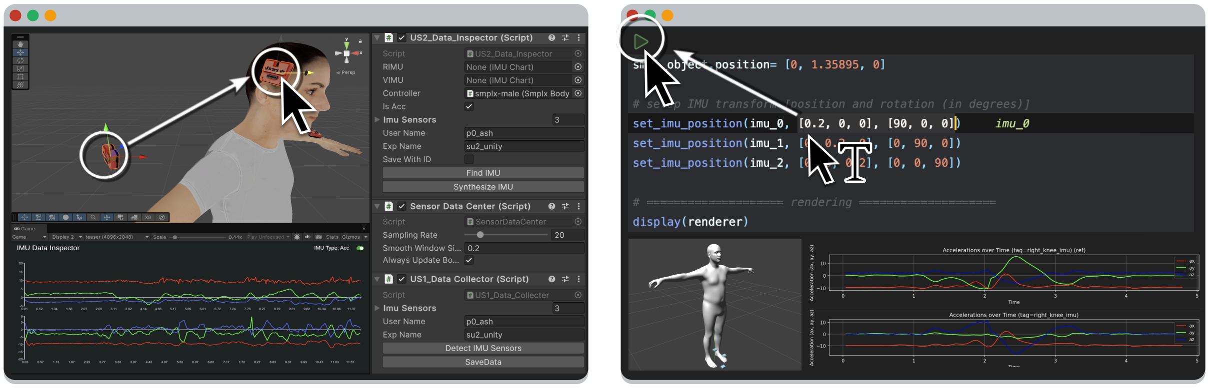
5.3 Evaluation Platforms and Apparatus

To evaluate Vsens, we designed a baseline workflow that reflects representative virtual IMU configuration paradigms: a 3D editor (Unity) [20, 61] and a code-based interface (e.g., Python) [71]. These platforms were selected to capture non-immersive, conventional practice used by HAR developers today. Each baseline was implemented with encapsulation layers that provide code templates, simplified interaction logic, and restrict the editing scope to focus participants on core configuration tasks. In addition to oral briefing, printed operation guides were provided for each platform to ensure that all participants could complete the tasks within a reasonable time and effort envelope. Additionally, it's necessary to note that rather than isolating the XR modality itself, our comparison aimed to evaluate the holistic workflow design that XR enables — spatial manipulation, embodied feedback, and real-time coupling between configuration and signal inspection. While prior desktop-based tools offer visualization functions, they typically operate in fragmented pipelines that separate configuration, signal synthesis, and analysis.

Baselines. For the *Unity* baseline, as shown in Figure 6.1, consisted of a pre-configured scene containing a digital avatar and a set of virtual IMU objects linked via C# scripts. Participants could select objects, adjust spatial parameters using transform gizmos (i.e., position and rotation manipulators), and trigger signal synthesis through the Inspector panel. Interaction was performed via mouse and keyboard. The *Python* baseline was implemented using PyCharm with pre-configured Jupyter notebooks that allowed participants to specify sensor positions and orientations by modifying parameter values and to export synthesized data as CSV files (Figure 6.2).

To ensure a fair comparison, our baseline setups in Unity and Python also supported immediate visual previews, such as 3D placement indicators and on-demand waveform plots, providing functionality comparable to the visualization aids available in existing virtual IMU workflows.

Apparatus. We implemented *Vsens* in Unity using Meta XR All-in-One SDK [19]. All platforms ran on Alienware m15 R4 laptop with Intel Core i7-10870H and NVIDIA GeForce RTX 3070 8GB GDDR6. The Meta Quest 3 headset was operated in pass-through mode and tethered to the laptop via a high-speed USB link. Hand held controllers are used for 3D interactions to ensure precision and tracking stability. The virtual avatar was based on the SMPL-X model with 54 joints [22], and all motion clips were sourced from BML-MoVi dataset with a sample rate of 60 fps [14].



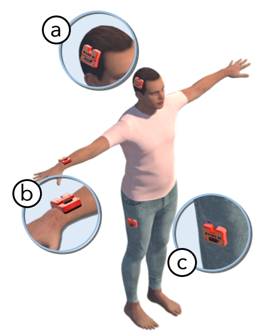
1. View of Unity Editor 2. View of Pycharm Editor

Fig. 6. Evaluation platforms (Unity and Python editor) used in the user study.

5.4 Tasks and Experimental Conditions

To evaluate the effectiveness of *Vsens* in supporting virtual IMU data collection, we designed a three-task study. Figure 7.1 summarizes each task's objective, related hypotheses, and evaluation platforms. Figure 7.2 visualizes the three virtual IMU locations used across tasks: *right-side of the head* (headphones), *right wrist* (smartwatch), and *right thigh* (phone pocket), corresponding to the representative configuration that is commonly seen in HAR systems [44, 69]. Detailed procedures and platform-specific actions are described in the following subsections.

Tasks	Hypotheses	Objective	Evaluation Platforms
Task 1: Sensor Placement	H1	Evaluate accuracy and efficiency of virtual IMUs placement.	Vsens v.s. Unity v.s. Python
Task 2: Signal Matching	H2	Evaluate how real-time feedback support the signal mapping.	Vsens v.s. Unity
Task 3: Positioning *Vsens* in Practical Workflows	H3	Explore how Vsens supports virtual IMU data collection in practical scenario.	-

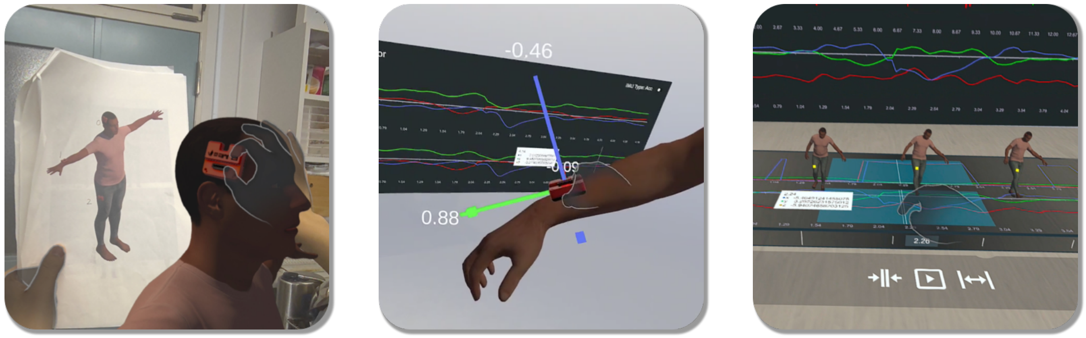


1. Objectives and Evaluation Platforms for each Task 2. Reference Virtual IMU Positions

Fig. 7. 1. The table contains the Tasks performed in the study, objectives of the tasks, and the evaluation platforms involved in each task; 2. Virtual IMU positions were used during the study.

5.4.1 *Task 1: Sensor Placement (H1).*

To validate H1, this task evaluated how accurately and efficiently participants could place virtual IMUs on a digital avatar to match a designated configuration, including both position and orientation. As shown in Figure 7.2, each participant was provided with a printed tri-view schematic of the target sensor layout and asked to replicate the configuration using different methods (*Vsens*, Unity, and Python). This task simulates the common practice in virtual IMU workflows of estimating placement based on video or



Task 1. Sensor Placement Task 2. Signal Matching Task 3. Practical Workflow

Fig. 8. A preview of participants' operation in Task 1, Task 2, and Task 3 during the user study.

annotated images. Participants performed the task using three methods in a counterbalanced order to mitigate learning effects. Each platform was preceded by a short practice trial to ensure familiarity with the interface.

As shown in Figure 8.1, participants can directly manipulate the virtual sensor with their hands in Vsens. While in Unity and Python, participants use the transform gizmos and keyboard/mouse input as described in Section 5.3. Participants were instructed to indicate when they were satisfied with the placement, at which point the trial was marked as complete.

5.4.2 Task 2: Signal Matching (H2). The goal of this task is to examine how users perform signal matching using *Vsens* versus a conventional Unity workflow. This task simulates a common scenario in HAR development where the exact sensor orientation is unknown and must be inferred by comparing synthesized and real IMU signals. We selected Unity as the sole baseline for this task, as it supports interactive visualization and parameter adjustment comparable to *Vsens*. Code-based tools such as Python are inherently script-driven and do not emphasize real-time features, making them unsuitable for a fair comparison focused on dynamic feedback and iterative refinement.

Participants were shown a motion animation (right leg kick) from the BML-MoVi dataset [14], loaded onto a digital avatar (Main Actor). To avoid memorization, for each condition (Vsens, Unity), randomized rotations were applied to the initial IMU orientation, and distinct reference signals were assigned. In the Unity condition, participants adjusted rotation values in the Unity inspector, and visualized real-time signal alignment in the Game panel. In *Vsens*, orientation was adjusted using direct controller manipulation (Figure 8.2), with synchronized panels visualizing virtual and reference signals in real time. Each trial ended when the participant felt their signal sufficiently matched the reference.

5.4.3 Task 3: Evaluating Vsens in Practical Workflows (H3). This task simulated a realistic end-to-end workflow to evaluate how *Vsens* supports virtual IMU data generation and how the resulting data compares to real measurements. Participants first wore real IMU sensors (Movella DOT[2]) attached to the body locations shown in Figure 7.2. The real IMUs served a dual purpose: (1) providing reference signals for virtual data, and (2) serving as ground truth for subsequent accuracy validation.

Each participant performed three common motion exercises — *knee kick*, *lunge*, and *jumping jack* — each lasting about 10 seconds, while real IMU and video footage were recorded. Immediately afterward, the videos of

[2]https://www.movella.com/

the motions were processed using WHAM [56] to reconstruct a 3D animation, which takes up to 30 seconds. The reconstructed animation, along with the real IMU data, was then imported into *Vsens*.

Participants were instructed to use all system functionalities described in Section 4. They could place, configure, and inspect multiple virtual IMUs on the digital avatars, adjust playback, and compare synthesized signals against their own recorded references. This open-ended setup allowed participants to freely replicate and refine their virtual data using real-time feedback. After participants felt satisfied with the configuration and signal quality, they exported the resulting virtual IMU data, which were later analyzed to assess fidelity and accuracy relative to the real measurements.

5.5 Measurements

Our measurements were designed to align with the hypotheses formulated in Section 5.1, combining quantitative performance metrics, subjective ratings, and qualitative feedback.

Task Completion Time (Task 1, Task 2). Measured as the duration between the start of sensor placement and the participant's confirmation of completion, recorded in seconds.

Iteration Count (Task 1, Task 2). Derived from the frame-encoded video logs to capture the number of meaningful repositioning or rotation actions. Two researchers independently annotated all recordings using a consistent coding scheme, resolving discrepancies through discussion.

NASA-TLX Scores (Task 1, Task 2). After completing both tasks, participants filled out NASA-TLX questionnaires to assess their perceived workload across conditions, providing a subjective complement to objective task metrics.

Placement Error (Task 1). Decomposed into positional and rotational components. Positional error was computed as the Euclidean distance between the participant's final placement and the designated reference point. Rotational error was calculated by converting normalized Euler angles ($[-180°, 180°)$) into quaternions and computing their angular difference from the reference orientation.

Signal Deviation (Task 2). Quantified using the dynamic time warping (DTW) distance between the synthesized and reference acceleration signals along X, Y, and Z axes. The average of the three axis-wise DTW scores was used as an overall deviation metric.

Classification Accuracy (Task 3). To evaluate the practical validity of virtual IMU data, we compared signals generated in *Vsens* with real IMU recordings of the same motions. All participants' datasets were aggregated, and a leave-one-out evaluation was conducted across three activity classes. The raw acceleration data were segmented using a 2-second sliding window (50% overlap) and smoothed with a moving-average filter ($k = 5$) to reduce noise. For traditional machine-learning models (Random Forest, Support Vector Classifier, Decision Tree), we used ECDF features [15] to represent temporal distributions. For the deep-learning baseline (ConvLSTM [47]), both virtual and real signals were resampled to 20 Hz and fed as temporal sequences. Each classifier was evaluated ten times with different random seeds, and the average accuracy was reported.

System Usability Scale (SUS) Scores (Post-study). After Task 3, participants completed the SUS questionnaire to rate the overall perceived usability of *Vsens* relative to conventional virtual IMU workflows.

Semi-structured Interview (Post-study). A post-study interview gathered qualitative reflections on usability, learning curve, and potential applications. Thematic analysis of these interviews is detailed in Section 5.7.

5.6 Quantitative Results

5.6.1 Data Analysis. For Tasks 1 and 2, we conducted statistical analysis on participants' performance metrics, including task completion time, iteration count, placement error, and signal deviation. Given the small sample size ($N = 20$), within-subjects design, and the non-normal distribution of most measures (assessed using the Shapiro-Wilk test), we applied non-parametric tests: the *Friedman test* for comparing three conditions in Task 1, and the *Wilcoxon signed-rank test* for the paired conditions in Task 2 [2]. When appropriate, post-hoc pairwise comparisons were conducted using Holm correction to control for multiple comparisons. We report statistically significant results using the following notation: $p < .05$ ($*$), $p < .01$ ($**$), and $p < .001$ ($***$). Descriptive statistics are presented as means (M) and standard errors (SE) for readability and ease of comparison, despite the use of non-parametric tests for inferential analysis.

During the study, all participants successfully completed all three tasks without exceeding the allotted time limits or reporting discomfort. No trial had to be excluded due to incomplete data or task failure.

5.6.2 Results of Task 1: Sensor Placement. We compared participant performance across three metrics: task completion time, iteration count, and placement error. Figure 9 shows boxplots of these measures across the three conditions (Vsens, Unity, and Python). A Friedman test revealed significant main effects of condition on both task completion time, $\chi^2(N{=}20) = 40.00, p < .001$, and iteration count, $\chi^2(N{=}20) = 40.00, p < .001$. Post-hoc pairwise comparisons with Holm correction indicated that participants completed the task significantly faster and with fewer iterations in *Vsens* (Time: $M = 30.33$ sec, $SE = 3.14$; Iterations: $M = 3.83$, $SE = 0.53$) compared to both Unity (Time: $M = 83.27$ sec, $SE = 7.80$; Iterations: $M = 9.90$, $SE = 1.09$) and Python (Time: $M = 154.10$ sec, $SE = 13.64$; Iterations: $M = 15.30$, $SE = 1.29$). These results demonstrate that *Vsens* substantially reduced both the time and number of adjustment iterations required for virtual IMU placement compared to conventional workflows, indicating improved efficiency and controllability during configuration.

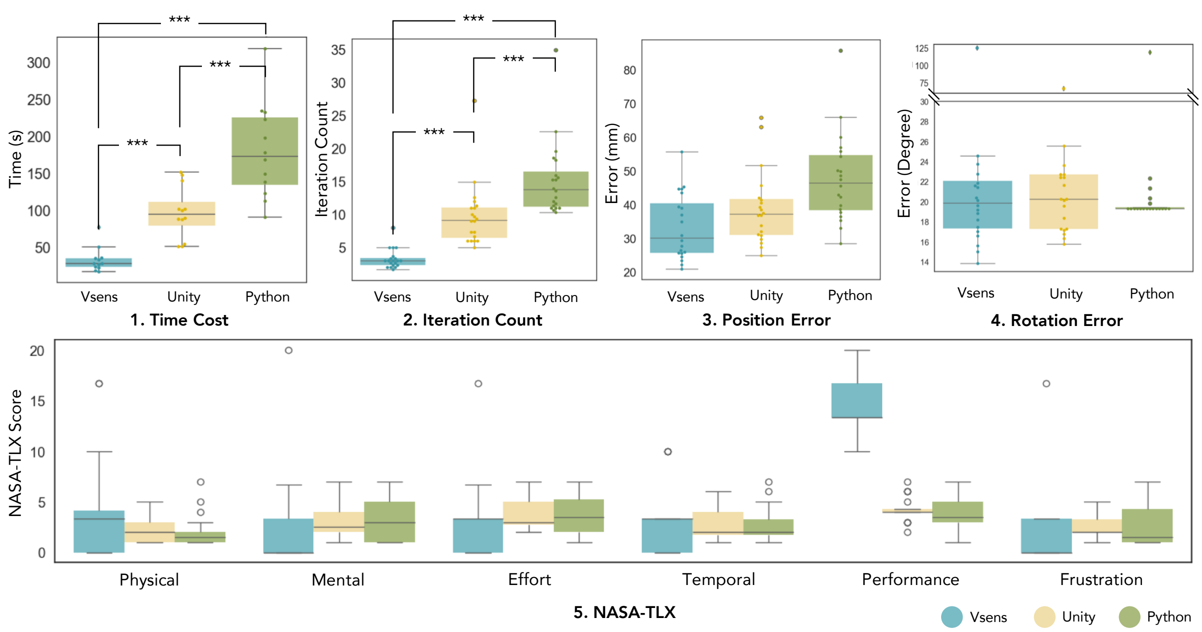


Fig. 9. Results from Task 1. Metrics including: 1. Time Cost; 2. Iteration Count; 3. Placement Error (Position); 4. Placement Error (Rotation); 5. NASA-TLX Score.

Despite improvements in task efficiency, placement accuracy showed no statistically significant differences across the three conditions for either positional or rotational error (Friedman test, $p > .05$). Positional and rotational deviations remained small and consistent across conditions, indicating comparable spatial precision regardless of workflow. This suggests that while *Vsens* accelerated the configuration process, it did not compromise spatial accuracy or induce the common usability-accuracy trade-off often observed in interactive configuration tasks. Thus, **H1** was supported.

Overall NASA-TLX scores reflected this pattern: participants reported slightly higher physical demand in *Vsens* due to the embodied operations, but markedly lower mental demand, effort, and frustration compared with the desktop conditions. They also felt less temporal pressure, suggesting that the immersive environment allowed a more natural and less stressful workflow. Interestingly, participants perceived considerably higher performance (precision) when using Vsens, even though objective accuracy was comparable across platforms. This highlights a perception–performance gap, where spatial feedback may have enhanced user confidence during manipulation. Another notable observation was the distribution of rotational errors shown in Figure 9.4. We identified a subset of participants who largely neglected virtual IMU orientation adjustment, resulting in outlier errors. Additionally, in the Python condition, participants consistently exhibited coarse adjustment behaviors, tending to modify rotation values in large steps (e.g., multiples of 90°) rather than fine-tuning. This led to similar final orientation errors across participants, but often diverging from the reference orientations.

5.6.3 Results of Task 2: Signal Matching. In Task 2, participants were asked to adjust the orientation of virtual IMUs so that the synthesized acceleration signals aligned with a given reference signal. Figure 10 summarizes the performance across the two conditions (*Vsens* vs. Unity). A Wilcoxon signed-rank test revealed that participants completed the orientation calibration significantly faster with *Vsens* ($W = 1.0$, $p < .001$). The mean task completion time for *Vsens* was $M = 107.26$ seconds ($SE = 13.89$), compared to $M = 263.38$ seconds ($SE = 28.36$) for Unity.

Iteration count showed a similar trend. Participants required significantly fewer adjustments in *Vsens* ($W = 29.50$, $p < .01$; $M = 4.75$, $SE = 0.49$) than in *Unity* ($M = 7.25$, $SE = 0.79$). These results indicate that the real-time feedback provided by *Vsens* substantially improved the efficiency of signal matching, allowing participants to converge toward alignment faster and with fewer iterative corrections.

For signal alignment accuracy, measured by the DTW distance between virtual and reference acceleration signals, participants achieved significantly lower deviations when using *Vsens* ($W = 5.0$, $p < .001$). This indicates that the real-time feedback and spatial interaction mechanisms in *Vsens* facilitated more precise orientation calibration and faster convergence toward optimal signal alignment.

Participants' subjective workload ratings from the NASA-TLX further corroborated these findings (Figure 10.4). Although perceived physical demand was comparable across conditions, participants reported markedly lower mental demand, effort, temporal pressure, and frustration when using *Vsens*. They also rated their perceived performance higher, suggesting greater confidence and a sense of control during the calibration process. Together, these results support **H2**, demonstrating that real-time feedback and spatially coupled visualization in *Vsens* not only improved objective task efficiency and accuracy but also reduced cognitive workload and enhanced user confidence during orientation matching.

5.6.4 Results of Task 3: Classification Accuracy. Task 3 evaluated the representational validity of virtual IMU data generated through *Vsens* by comparing its classification accuracy with that of real IMU recordings collected in the same motion exercises. As shown in Figure 11, we trained four representative classifiers (RF, SVC, DT) and ConvLSTM on both datasets to recognize three activities (*knee kick*, *lunge*, and *jumping jack*).

Overall, the results demonstrate that virtual IMU data produced using *Vsens* achieved comparable recognition performance to real IMU across most models, indicating **H3** is supported. For example, accuracies with RF and SVC exceeded 0.97 when combining all IMUs, closely matching the results obtained from real data. The relatively high accuracy across classifiers can also be attributed to the simplicity and distinct frequency patterns of the three selected motion exercises. These findings indicate that the virtual IMU data generated via *Vsens* are sufficiently reliable for typical HAR development and model prototyping. Our system does not aim to improve classification algorithms directly; instead, by enabling more controllable, interpretable, and efficient virtual

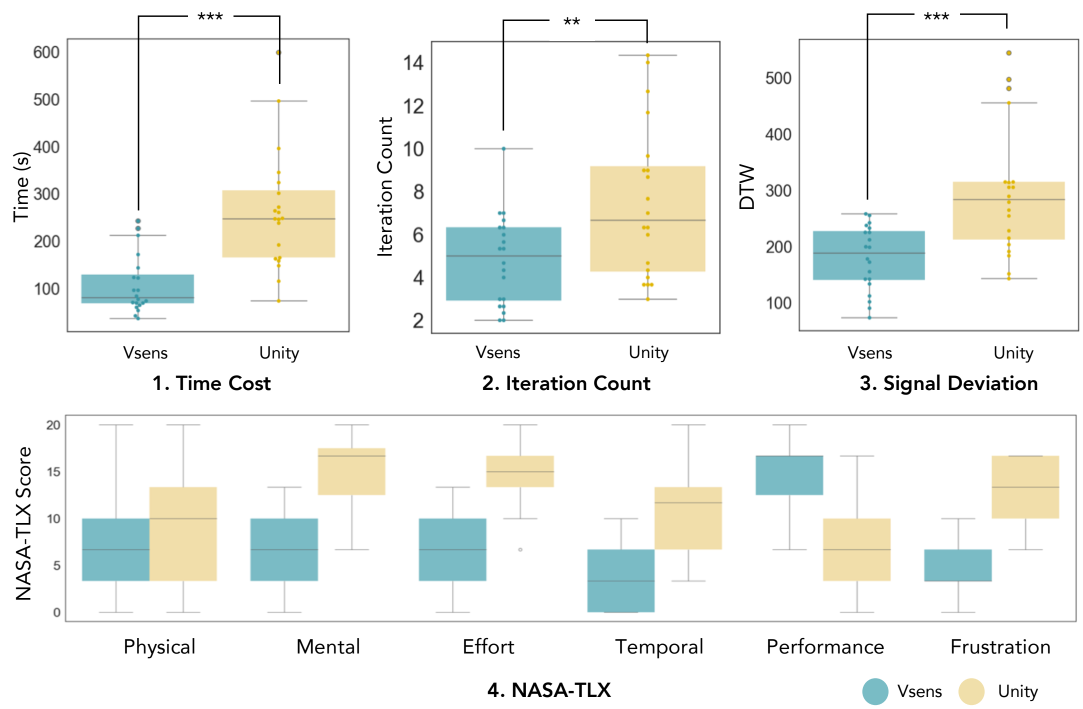


Fig. 10. Results from Task 2: 1. Time Cost; 2. Iteration Count; 3 Signal Deviation (DTW); 4. NASA-TLX Score.

	RF		SVC		DT		ConvLSTM	
	Virtual	Real	Virtual	Real	Virtual	Real	Virtual	Real
IMU (head)	0.591	**0.756**	0.574	**0.595**	0.538	**0.668**	0.520	**0.616**
IMU (wrist)	**0.896**	0.850	0.864	**0.889**	0.850	**0.864**	0.693	**0.734**
IMU (thigh)	0.776	**0.800**	0.691	**0.786**	0.696	**0.777**	0.609	**0.649**
All IMUs	0.984	**0.992**	**0.976**	0.964	**0.953**	0.895	0.725	**0.891**

Fig. 11. Accuracy of classifying three motion exercises using real IMU data and virtual IMU data generated by *Vsens*.

data authoring, it indirectly supports the development of effective HAR models. Extensive ongoing work has explored improving virtual IMU fidelity and classification accuracy through physics-based simulation and data augmentation techniques [18, 45, 67]; *Vsens* is complementary to these efforts, focusing on the human-in-the-loop interface that enables such techniques to be applied more effectively.

5.6.5 System Usability Scale (SUS) Evaluation. After completing all tasks, participants filled out the System Usability Scale (SUS) to provide an overall assessment of the *Vsens* workflow compared with the conventional Unity- and Python-based methods. Figure 12 summarizes the item-level responses. Overall, *Vsens* receives

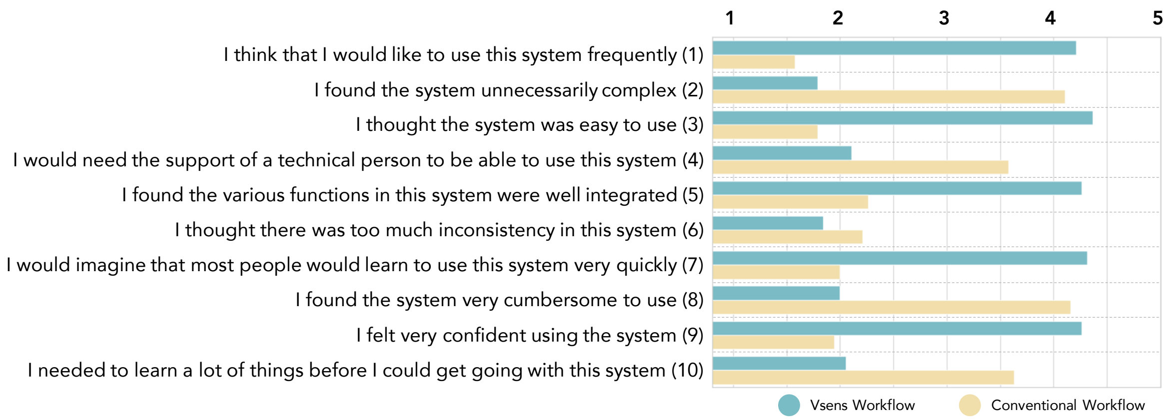


Fig. 12. Results of averaged SUS scores after all the tasks, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

consistently higher ratings on easier to use, integration of functions, and user confidence, while scoring lower on perceived complexity and cumbersomeness. Participants also felt they could learn and operate the system more quickly, suggesting a smoother learning curve and greater sense of control. These tendencies collectively indicate a clear user preference for the integrated, feedback-driven workflow provided by *Vsens*. We report the absolute SUS scores here to provide additional context alongside the relative comparison between systems. The proposed system achieved a mean SUS score of 60.9 (SD = 26.7), compared to 53.2 (SD = 21.5) for the conventional workflow.

5.7 Qualitative Results

The post-study interviews further contextualized these findings. We conducted semi-structured interviews with all 20 participants. Interviews were conducted in various languages, and the recordings were translated into English after speech-to-text. Transcripts were analyzed using inductive thematic analysis by two researchers [46]: first open-coded by two researchers independently, followed by iterative clustering into higher-order themes. Coding disagreements were discussed and refined through consensus. After the analysis, 3 recurring themes emerged: (1) *Embodied Interaction and Usability*, (2) *Real-Time Feedback and Visualization*, (3) *Exploratory Use and Broader Applications*, (4) *Perceptions of Virtual Data Accuracy and Trustworthiness.*

5.7.1 Theme 1: Embodied Interaction and Usability. Participants consistently described the embodied nature of interaction in *Vsens* as a key factor that made sensor configuration more intuitive and spatially comprehensible. Compared with the mouse and keyboard operations in conventional workflows, participants emphasize that being able to manipulate position and orientation simultaneously mirrors how they naturally think about IMU placement in reality, for example:

> "*I can modify both position and orientation simultaneously, which in Unity requires six separate values.*" (PA-10), and "*When I grab and rotate the sensor directly in 3D, I immediately understand which axis I'm changing — there's no mental translation needed.*" (PA-7).

Participants further noted that the immersive environment and axis visualization enhanced the inspection and depth perception. Such spatial awareness made fine-tuning "*feel like sculpting*" rather than numerical adjustment (PA-1), which resemble to physical practice as described:

> “*To check if it (IMU) is really on the skin, I just move my head closer. In Unity, you’d rotate the whole scene.*” (PA-4), and "*The visualization of the axis help me understand how IMU is oriented easily.*" (PI-1).

Beyond intuitiveness, several participants — particularly those from clinical and biomechanics backgrounds (PCs) — emphasized that the embodied interface substantially lowered the technical entry barrier for non-engineering users with few programming or 3D editor experiences. Many clinicians admitted that traditional workflows felt "*daunting*" or "*too programming-heavy*" (PC-1, PC-4), whereas the XR-based interaction enabled them to focus on *what* rather than *how* of configuration, for example:

> "*I don’t know Unity or coding, but here I can just look at the avatar, grab the sensor, and see the signal right away. It’s much easier than I expected when you introduced it to me.*" (PC-1). Another rehabilitation practitioner echoed, "*... I felt I could configure virtual IMUs without asking an engineer to help.*" (PC-5).

These comments illustrate how embodied spatial interaction not only enhances the intuitiveness or efficiency of the virtual sensor manipulation by HAR developers in academia or industry, but also broadens usability for domain experts with limited computational experience. By aligning virtual sensor interface with bodily perception and spatial reasoning, *Vsens* bridges disciplinary divides between engineering-oriented and application-oriented HAR practitioners.

5.7.2 Theme 2: Real-Time Feedback and Visualization. Participants emphasized that the real-time and embodied feedback during Task 2 was one of the most valuable aspects of *Vsens.* The immediate response between physical movement and signal change allowed users to reason about orientation without performing explicit coordinate calculations or reasoning. Several participants noted that they no longer needed to interpret axis-wise acceleration values, as they could instead rely on visual trends, for example:

> “*In Unity, I considered the signs (of the acceleration) of each axis. I didn’t have to think about axes in Vsens. I just rotated the sensor with my hand while watching the signal change.*” (PA-4), and “*I looked at the data panel while rotating my hand. If the trend looked right, I kept going. If not, I’d reverse.*” (PA-3). When being asked to compare the real-time feedback of the Unity, one participant responded: "*...it’s a totally different feeling between dragging gizmos and rotating my wrist, which is much more intuitive.*" (PI-1).

This continuous coupling between motion and feedback reduced cognitive burden and helped participants build an intuitive sense of how IMU orientation affected signal patterns.

Beyond reducing effort, participants noted that the immediacy of feedback encouraged a more exploratory and less error-averse mindset. Because the visualized signals updated continuously, users felt comfortable experimenting and correcting mistakes on the fly rather than carefully planning each adjustment. Comments from participants suggest that real-time feedback not only accelerates alignment tasks but also reshapes the interaction into an iterative, confidence-building process where exploration becomes part of learning:

> “*The cost of making mistakes is much lower with real-time feedback. It doesn’t feel frustrating if I’m off.*” (PA-11), and “*It gives me a kind of positive reinforcement that helps me do it faster. Even with a slight delay, the updates were totally usable.*” (PA-8), or "*... so I can move the sensor along the (human) model surface to see what is changing.*" (PC-4).

Additionally, participants further highlighted the usefulness of visualization components. For example, the timeline animation and the configurable data panels are complementary supports to real-time feedback. Although the study tasks involved only short motion sequences, several participants appreciated the inclusion of an animation preview, describing it as a "*forward-looking feature*" that would be valuable for longer or more complex sequences (PI-2, PC-1, PC-3). A few participants suggested further refinements, such as allowing the entire interface — including the digital avatar and panels — to scale together so that the system could be used comfortably while seated (PA-8, PC-2). Overall, participants found the visualization design clear and accommodating, reinforcing the sense of control and fluidity that characterized interaction in *Vsens*.

5.7.3 Theme 3: Exploratory Use and Broader Applications. During the interview, participants often described *Vsens* as more than a data-collection tool, seeing it as a gateway for exploration in HAR development. Several participants noted that while they had heard of virtual sensors or digital-twin systems (PA-1, PA-7, PC-1), this was their first time directly experiencing one, which lowered the entry barrier, making the idea of virtual sensing "*approachable and practical*", for example:

> "*I knew what virtual sensors were, but never actually used them. This system made it feel worth trying.*" (PA-7)

Faculty members and researchers envisioned broader uses in education and prototyping. They suggested that *Vsens* could serve as a teaching aid to illustrate complicated concepts such as sensor axes and coordinate transformations without requiring hardware setup, for example:

> "*Rotations and coordinate frames are hard for beginners to grasp; showing them in HMD makes it click.*" (PA-1), and "*This could be a potential way to demonstrate how the sensors work to students in medical school.*" (PC-2). Others emphasized its value for rapid experimentation, enabling quick deployment and repositioning of IMUs to test new sensing configurations, as one researcher noted: "*It lets me easily try unconventional placements, which is common in HCI, like NurseCare [13] or other HAR systems.*" (PA-7)

Clinical participants also suggested future possibilities beyond the current scope. They would like domain-specific modules that simulate pathological motions (e.g., valgus knees or elbow deviations), or include anatomical constraints for more precise modeling (PCs). While acknowledging that these features extend beyond the present focus, they regarded them as meaningful directions for applying virtual sensing as the interface to healthcare and accessibility [35, 36, 55], and *Vsens* made them "*more open to engaging with*" such emerging technologies (PC-1, PC-2, PC-4). Overall, these reflections illustrate how *Vsens* not only streamlines current virtual sensor workflows but also inspires new practices and applications across contexts.

5.7.4 Theme 4: Perceptions of Virtual Data Accuracy and Trustworthiness. Although the study primarily focused on interaction and workflow design, nearly all participants spontaneously raised questions about the general reliability of virtual IMU data. Their comments revealed that concerns about data fidelity remain central to how users conceptualize virtual sensing, regardless of their background.

Researchers and engineers generally recognized that virtual IMUs cannot fully replicate real-world conditions due to factors such as deformation, contact dynamics, and domain shift. Yet they viewed synthesized data as effective for early-stage prototyping and exploratory testing, where relative patterns and signal shapes matter more than absolute precision. Participants with development experience thus framed virtual sensing as a pragmatic trade-off between fidelity and cost, noting that recruiting participants and conducting physical pilot studies often impose significant logistical barriers, for example:

> "*Real sensors still feel more trustworthy, but the simulation will keep improving.*" (PA-3), "*It can't replicate my phone moving in my pocket, but that's okay — it gives a clean baseline for early testing.*" (PA-4), and *As a prototyping tool, even if the accuracy isn't perfect, it's good enough to make design decisions.*" (PA-8)

Clinical participants, on the contrary, expressed stronger reservations. Given their professional responsibility for diagnosis and patient monitoring, they regarded reliability as non-negotiable. One medical doctor said "*definitely no (with laughter)*" when commenting on the synthesized data quality (PC-1). Another explained, "*For* *our field, accuracy always comes first — even at early stages.*" (PC-5). Some nonetheless acknowledged that in less critical contexts, such as general movement guidance or rehabilitation training [65], virtual data might serve as a useful reference. As PC-1 summarized, "*I'm very excited about this technology — we can do so much — but realistically, it will take years before the virtual data becomes truly reliable.*".

Overall, these perspectives underscore a shared optimism tempered by professional caution: while engineers and researchers emphasize feasibility and iteration, clinicians foreground reliability and accountability. This divergence suggests that future virtual sensing systems need to balance convenience and precision differently across application domains.

6 DISCUSSION

6.1 Revisiting Challenges

This section revisits the three challenges identified in Section 3 and consolidates the overall connections throughout the paper: the initial *challenge framing*, system *design rationales*, *hypotheses*, experimental *tasks*, and *resulting evidences* into a coherent narrative. Each challenge is discussed in light of how *Vsens* addressed it through its integrated workflow and feedback mechanisms, and how HAR developers from diverse backgrounds responded to these design decisions.

6.1.1 Revisit Challenge 1: Complex Interaction. This challenge concerned the difficulty of configuring and inspecting virtual IMUs across disjointed tools. *Vsens* was designed to address this issue through spatial manipulation, visualization, and data synthesis within an embodied XR environment. Quantitative results in Task1 and 2 (Section 5.6) confirm that participants completed placement tasks significantly faster while maintaining comparable or even improved precision and reporting lower workload on the NASA-TLX scale. Beyond these quantitative gains, SUS scores and interview feedback (Section 5.7) further support these findings: participants described the conventional workflows as "*tedious*," "*hard to keep track of*", and "*like coding blindfolded*". These reflections illustrate that the pain points summarized in Section 3 are indeed experienced by HAR developers in practice — regardless of background or expertise — and that the proposed workflow can substantially mitigate the complexity of traditional virtual-sensor configuration.

6.1.2 Revisit Challenge 2: Lack of Real-time Inspection and Feedback. While conventional tools, including our encapsulated Unity baseline, can display immediate previews, such feedback remains functionally detached from spatial manipulation and often confined to a secondary panel. Users must still alternate between adjusting parameters and cognitively mapping those changes to waveform variations, which disrupts the sense of continuity during configuration. What distinguishes *Vsens* is not only immediacy of feedback, but its spatial and embodied co-location—feedback that can be felt through action, not merely seen on a screen. Every movement of a virtual sensor is reflected immediately in the co-located signal view, allowing developers to interpret signal changes through direct spatial action rather than abstract value editing. Task 2 simulated this context through the signal-matching scenario, where participants aligned virtual IMU signals with reference data. Although both workflows offered instantaneous previews, participants using *Vsens* completed orientation calibration significantly faster and achieved lower signal deviation. Interview comments attributed this improvement to the tighter spatial coupling between manipulation and response — participants described the process as "*seeing the waveform move with my hand*" and "*finally understanding what each axis meant*" (Section 5.7). These results suggest that the benefit of *Vsens* lies not in providing feedback per se, but in transforming feedback into an integral part of reasoning and control, turning configuration into an interactive, interpretable process rather than a sequence of disjointed adjustments.

6.1.3 Revisit Challenge 3: Reliable Virtual Data. Ensuring that synthesized IMU signals faithfully approximate real-world measurements remains an open challenge that extends beyond interface design. Variations in motion dynamics, contact modeling, and domain shifts can all introduce discrepancies that no single simulation framework can fully resolve. Our intent with *Vsens* was therefore not to eliminate this gap, but to provide HAR developers with a more transparent and controllable workflow for examining and refining virtual data quality. Task 3 served as an initial demonstration of this capability. By comparing virtual and real IMU recordings in a simple three-class recognition task, participants were able to confirm that the signals generated in *Vsens* produced comparable accuracy across standard classifiers (Section 5.6.4). While these results cannot generalize to complex or large-scale deployments, they suggest that *Vsens* can already serve as a practical platform for benchmarking and early-stage model prototyping. Looking forward, we envision that coupling such authoring environments with domain-adaptation techniques, biomechanical modeling, or physics-informed learning could further bridge the remaining realism gap. In this sense, *Vsens* could serve as a reference for exploring how reliable virtual data generation can become an interactive and iterative process rather than a purely algorithmic one.

Taken together, these reflections illustrate that the core difficulties identified in virtual IMU workflows are not isolated technical issues but manifestations of fragmented and delayed feedback in the process. Our comparative evaluation with a diverse group of HAR developers—from academic, industrial, and clinical backgrounds—empirically substantiated the relevance of these challenges and confirmed that the proposed design goals align with the real needs of this community, demonstrating how addressing them in concert can lead to more transparent and human-centered approaches to virtual sensing.

6.2 Towards More Approachable and Usable Virtual Sensing Workflows

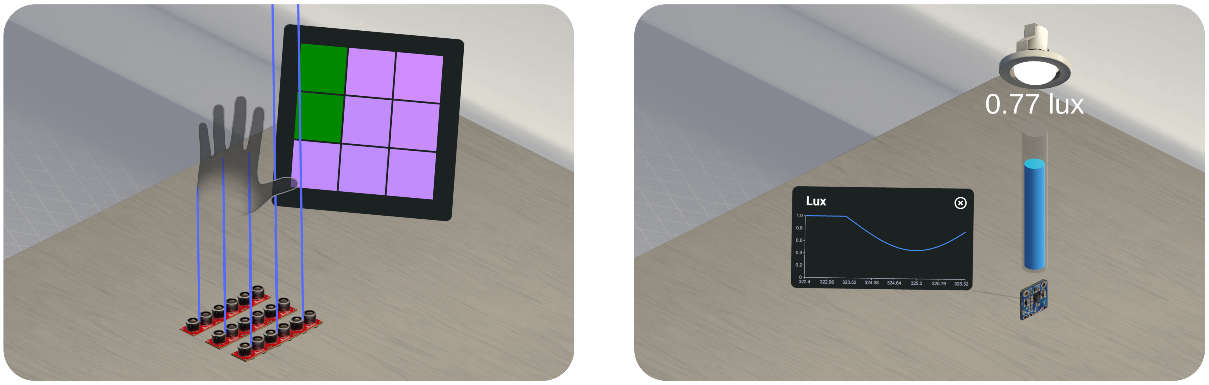
HAR systems form a foundational layer of ubiquitous computing research. Within this ecosystem, virtual sensing systems serve as powerful tools that assist in developing, validating, and refining HAR algorithms. However, the individuals who contribute to HAR system development are highly diverse in background and expertise, which cannot be easily categorized as "*experts*" or "*novices*"; instead, they represent different domains of practice. For example, a medical doctor studying biomechanics may be an expert in motion analysis but unfamiliar with computer science; an acoustic-sensing researcher may not have prior experience with inertial sensors; and even developers accustomed to IMU-based systems may lack proficiency in 3D editing environments.

Our participant pool (Section 5.2) was designed to reflect this diversity. Rather than limiting recruitment to the small community already experienced with virtual sensing, most of the participants are practitioners from academic, industrial, and clinical settings who had not previously used virtual IMUs. This decision aligns with the system's goal of lowering the barrier of entry and making virtual sensing technologies more usable and interpretable for developers across different subdomains. Within the study sessions, all participants were able to produce valid virtual IMU datasets with *Vsens*, and the classification results (Section 5.6.4) demonstrate that these datasets are practically usable. This outcome illustrates that *Vsens* can support both HAR developers with various backgrounds in authoring virtual sensor data without requiring deep prior knowledge of simulation tools. By improving the usability and approachability of virtual sensing, *Vsens* aims to indirectly advance HAR system development, enabling more researchers and practitioners to engage with sensor data creation as part of an integrated design process.

While *Vsens* currently focuses on the IMU modality as a concrete example, we envision the system could be conceived as a template or extensible interface for broader virtual sensing research. Future iterations may incorporate additional sensing modalities or simulations, enabling developers to compose multi-modal simulations that better capture the richness of real-world HAR contexts.

6.3 Beyond Virtual IMU Sensors

While *Vsens* was developed with a primary focus on virtual IMU data collection for HAR tasks, the underlying design motivations extend beyond this specific sensing modality. IMU-based systems are both ubiquitous and representative of sensing applications that involve multi-axis, high-frequency dynamics—conditions under which embodied interaction, spatial manipulation, and real-time feedback are most critical to usability and understanding.



1. Distance Sensor 2. Optical Sensor

Fig. 13. Virtual distance and optical sensor can be implemented in Vsens.

At the same time, the design principles demonstrated by *Vsens* — feedback coupling, spatial manipulation, and multi-modal visualization — are not unique to inertial sensing. They can be generalized to other sensor types such as proximity, optical, acoustic, where spatial configuration and signal interpretation play similarly central roles (Figure 13). For instance, in smart environment prototyping, XR interfaces could enable developers to position virtual environmental sensors within a reconstructed 3D model of a room [60], simulate their fields of detection, and iteratively adjust coverage or sensitivity based on synthetic output. Such workflows could streamline early-stage testing and reduce the material and temporal cost of building ubiquitous computing systems [16, 26].

Nevertheless, the current state of XR and simulation technologies imposes clear practical limits. Accurately replicating real-world physics, lighting, and acoustic propagation in real time remains challenging, and constructing faithful digital twins of user environments is often infeasible outside controlled settings. As a result, XR-based virtual sensing should be viewed not as a replacement for physical experimentation but as a complementary medium for conceptualization and pre-deployment reasoning.

Looking forward, as XR rendering fidelity and physical simulation frameworks mature, we anticipate gradual convergence between virtual and physical sensing workflows [16]. Rather than promising full environmental replication, future systems may focus on enabling partial, interactive, and explainable simulations that help developers explore sensing concepts before real-world implementation.

7 LIMITATIONS

While the study demonstrates the effectiveness and potential of *Vsens*, several limitations should be acknowledged and addressed in future work.

Scope of system integration. Although *Vsens* supports end-to-end virtual data collection, the current prototype does not yet integrate downstream processes such as annotation, dataset management, or model training. Extending the system to include these components could enable a more seamless pipeline—from data generation to model evaluation—and provide a deeper understanding of how virtual data interacts with machine learning workflows.

Study design and identified challenges. The study defined its core challenges based on prior literature and validated them through a comparative user study rather than a separate formative investigation. This approach provided both quantitative and qualitative evidence of *Vsens*'s effectiveness but may still carry blind spots regarding issues specific to existing virtual sensing practices. Future research could complement this comparative approach with formative or ethnographic studies to further contextualize HAR developer needs regarding using virtual sensing technology in real-world settings.

Participant background and representativeness. There were four of the twenty HAR developers had prior experience with virtual sensing systems. This sampling choice reflected the current scarcity of such specialists and our intention to include a diverse range of HAR developers, including those without prior knowledge. Nevertheless, with a larger population of experienced virtual-sensing researchers, future studies could yield more domain-specific insights and richer comparative perspectives on workflow adoption and technical depth.

Task specificity and study environment. Our evaluation was conducted through controlled, task-based sessions within a laboratory setting. While this allowed for systematic measurement and repeatability, it does not fully capture how developers might use *Vsens* in long-term, open-ended prototyping or during actual dataset creation. Future in-the-wild deployments and longitudinal studies could examine how the system supports creativity, iteration, and potential collaboration over extended periods.

8 CONCLUSION

This paper presented *Vsens*, an XR-based system that integrates spatial interaction, real-time feedback, and visualization into a unified and embodied workflow for virtual IMU data collection. Through a controlled study with 20 HAR developers from diverse disciplinary backgrounds, we showed that the system improves configuration efficiency and interpretability compared with conventional virtual sensing workflows. The results inform future development of more usable and approachable virtual sensing environments that emphasize transparency and ease of iteration. By treating virtual data generation as an interactive and interpretable process, rather than a purely automated one, we aim to indirectly support the broader process of HAR system development.

Acknowledgments

This work was partially supported by JST Keio-SPRING (Grant Number Y01GQ25172, Y01GQ25160), JST AIP Acceleration Research (Grant Number JPMJCR25U4), and JSPS KAKENHI (Grant Number JP25K03168). Part of this work was supported by the Natural Science Basic Research Program of Shaanxi (Program No.2025JC-YBQN-852) and China Postdoctoral Science Foundation (No.2024M762548). We'd also like to thank Yixuan Zhang for sketching the figures in this paper.

References

[1] Karan Ahuja, Yue Jiang, Mayank Goel, and Chris Harrison. 2021. Vid2Doppler: Synthesizing Doppler Radar Data from Videos for Training Privacy-Preserving Activity Recognition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 292, 10 pages. doi:10.1145/3411764.3445138

[2] Milton Friedman and. 1937. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Amer. Statist. Assoc.* 32, 200 (1937), 675–701. doi:10.1080/01621459.1937.10503522 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1937.10503522

[3] Oresti Banos, Mate Attila Toth, Miguel Damas, Hector Pomares, and Ignacio Rojas. 2014. Dealing with the Effects of Sensor Displacement in Wearable Activity Recognition. *Sensors* 14, 6 (2014), 9995–10023. doi:10.3390/s140609995

[4] Sejal Bhalla, Mayank Goel, and Rushil Khurana. 2022. IMU2Doppler: Cross-Modal Domain Adaptation for Doppler-based Activity Recognition Using IMU Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 145 (Dec. 2022), 20 pages. doi:10.1145/3494994

[5] Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. 2021. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 470, 15 pages. doi:10.1145/3411764.3445651
[6] Yize Cai, Baoshen Guo, Flora Salim, and Zhiqing Hong. 2025. Towards Generalizable Human Activity Recognition: A Survey. arXiv:2508.12213 [eess.SP] https://arxiv.org/abs/2508.12213
[7] Michelle Carney, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. 2020. Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3334480.3382839
[8] Ciprian Chelba, Tomás Mikolov, Mike Schuster, Qi Ge, Thorsten Brants, and Phillipp Koehn. 2013. One Billion Word Benchmark for Measuring Progress in Statistical Language Modeling. *CoRR* abs/1312.3005 (2013). arXiv:1312.3005 http://arxiv.org/abs/1312.3005
[9] Wenqiang Chen, Shupei Lin, Elizabeth Thompson, and John Stankovic. 2021. SenseCollect: We Need Efficient Ways to Collect On-body Sensor-based Human Activity Data! *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 91 (Sept. 2021), 27 pages. doi:10.1145/3478119
[10] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255. doi:10.1109/CVPR.2009.5206848
[11] Adrian Derungs, Julia Seiter, Corina Schuster-Amft, and Oliver Amft. 2015. Estimating physical ability of stroke patients without specific tests. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (Osaka, Japan) *(ISWC '15)*. Association for Computing Machinery, New York, NY, USA, 137–140. doi:10.1145/2802083.2808412
[12] Kayla DesPortes, Aditya Anupam, Neeti Pathak, and Betsy DiSalvo. 2016. BitBlox: A Redesign of the Breadboard. In *Proceedings of the The 15th International Conference on Interaction Design and Children* (Manchester, United Kingdom) *(IDC '16)*. Association for Computing Machinery, New York, NY, USA, 255–261. doi:10.1145/2930674.2930708
[13] Maximilian Dürr, Carla Gröschel, Ulrike Pfeil, and Harald Reiterer. 2020. NurseCare: Design and 'In-The-Wild' Evaluation of a Mobile System to Promote the Ergonomic Transfer of Patients. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376851
[14] Saeed Ghorbani, Kimia Mahdaviani, Anne Thaler, Konrad Kording, Douglas James Cook, Gunnar Blohm, and Nikolaus F. Troje. 2021. MoVi: A large multi-purpose human motion and video dataset. *PLOS ONE* 16, 6 (06 2021), 1–15. doi:10.1371/journal.pone.0253157
[15] Nils Y. Hammerla, Reuben Kirkham, Peter Andras, and Thomas Ploetz. 2013. On preserving statistical characteristics of accelerometry data using their empirical cumulative distribution. In *Proceedings of the 2013 International Symposium on Wearable Computers* (Zurich, Switzerland) *(ISWC '13)*. Association for Computing Machinery, New York, NY, USA, 65–68. doi:10.1145/2493988.2494353
[16] Seung Hyeon Han, Yeeun Han, Kyeongho Park, Sangjun Lee, and Woohun Lee. 2025. SpatIO: Spatial Physical Computing Toolkit Based on Extended Reality. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 978, 22 pages. doi:10.1145/3706598.3713747
[17] Jonggi Hong, Kyungjun Lee, June Xu, and Hernisa Kacorri. 2020. Crowdsourcing the Perception of Machine Teaching. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376428
[18] Zhiqing Hong, Zelong Li, Shuxin Zhong, Wenjun Lyu, Haotian Wang, Yi Ding, Tian He, and Desheng Zhang. 2024. CrossHAR: Generalizing Cross-dataset Human Activity Recognition via Hierarchical Self-Supervised Pretraining. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 2, Article 64 (May 2024), 26 pages. doi:10.1145/3659597
[19] Meta Horizon. 2025. https://developers.meta.com/horizon/downloads/package/oculus-platform-sdk/
[20] Lingtao Huang and Chengshuo Xia. 2024. ModifyAug: Data Augmentation for Virtual IMU Signal based on 3D Motion Modification Used for Real Activity Recognition. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 243, 7 pages. doi:10.1145/3613905.3650806
[21] Leslie Ikemoto, Okan Arikan, and David Forsyth. 2009. Generalizing motion edits with Gaussian processes. *ACM Trans. Graph.* 28, 1, Article 1 (feb 2009), 12 pages. doi:10.1145/1477926.1477927
[22] Max Planck Institute. [n. d.]. https://gitlab.tuebingen.mpg.de/jtesch/smplx-unity/-/tree/master/Assets/SMPLX
[23] Hye-Young Jo, Laurenz Seidel, Michel Pahud, Mike Sinclair, and Andrea Bianchi. 2023. FlowAR: How Different Augmented Reality Visualizations of Online Fitness Videos Support Flow for At-Home Yoga Exercises. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 469, 17 pages. doi:10.1145/3544548.3580897
[24] Jacob Kamminga, Rob van der Duim, and Erik ten Hove. 2024. M-MOVE-IT: Multimodal Machine Observation and Video-Enhanced Integration Tool for Data Annotation. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Melbourne VIC, Australia) *(UbiComp '24)*. Association for Computing Machinery, New York, NY, USA, 911–915. doi:10.1145/3675094.3678479

[25] Yoonji Kim, Hyein Lee, Ramkrishna Prasad, Seungwoo Je, Youngkyung Choi, Daniel Ashbrook, Ian Oakley, and Andrea Bianchi. 2020. SchemaBoard: Supporting Correct Assembly of Schematic Circuits using Dynamic In-Situ Visualization. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 987–998. doi:10.1145/3379337.3415887

[26] Yoonji Kim, Junyi Zhu, Mihir Trivedi, Dishita Turakhia, Ngai Hang Wu, Donghyeon Ko, Michael Wessely, and Stefanie Mueller. 2022. SensorViz: Visualizing Sensor Data Across Different Stages of Prototyping Interactive Objects. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 987–1001. doi:10.1145/3532106.3533481

[27] Orhan Konak, Robin van de Water, Valentin Döring, Tobias Fiedler, Lucas Liebe, Leander Masopust, Kirill Postnov, Franz Sauerwald, Felix Treykorn, Alexander Wischmann, Hristijan Gjoreski, Mitja Luštrek, and Bert Arnrich. 2023. HARE: Unifying the Human Activity Recognition Engineering Workflow. *Sensors* 23, 23 (2023). doi:10.3390/s23239571

[28] Hyeokhyen Kwon, Gregory D. Abowd, and Thomas Plötz. 2019. Handling annotation uncertainty in human activity recognition. In *Proceedings of the 2019 ACM International Symposium on Wearable Computers* (London, United Kingdom) *(ISWC '19)*. Association for Computing Machinery, New York, NY, USA, 109–117. doi:10.1145/3341163.3347744

[29] Hyeokhyen Kwon, Catherine Tong, Harish Haresamudram, Yan Gao, Gregory D. Abowd, Nicholas D. Lane, and Thomas Plötz. 2020. IMUTube: Automatic Extraction of Virtual on-body Accelerometry from Video for Human Activity Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 87 (Sept. 2020), 29 pages. doi:10.1145/3411841

[30] Hyeokhyen Kwon, Bingyao Wang, Gregory D. Abowd, and Thomas Plötz. 2021. Approaching the Real-World: Supporting Activity Recognition Training with Virtual IMU Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 111 (sep 2021), 32 pages. doi:10.1145/3478096

[31] Michael Laielli, James Smith, Giscard Biamby, Trevor Darrell, and Bjoern Hartmann. 2019. LabelAR: A Spatial Guidance Interface for Fast Computer Vision Image Collection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 987–998. doi:10.1145/3332165.3347927

[32] Jaewook Lee, Andrew D. Tjahjadi, Jiho Kim, Junpu Yu, Minji Park, Jiawen Zhang, Jon E. Froehlich, Yapeng Tian, and Yuhang Zhao. 2024. CookAR: Affordance Augmentations in Wearable AR to Support Kitchen Tool Interactions for People with Low Vision. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 141, 16 pages. doi:10.1145/3654777.3676449

[33] Woojin Lee, Ramkrishna Prasad, Seungwoo Je, Yoonji Kim, Ian Oakley, Daniel Ashbrook, and Andrea Bianchi. 2021. VirtualWire: Supporting Rapid Prototyping with Instant Reconfigurations of Wires in Breadboarded Circuits. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Salzburg, Austria) *(TEI '21)*. Association for Computing Machinery, New York, NY, USA, Article 4, 12 pages. doi:10.1145/3430524.3440623

[34] Zikang Leng, Amitrajit Bhattacharjee, Hrudhai Rajasekhar, Lizhe Zhang, Elizabeth Bruda, Hyeokhyen Kwon, and Thomas Plötz. 2024. IMUGPT 2.0: Language-Based Cross Modality Transfer for Sensor-Based Human Activity Recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 3, Article 112 (Sept. 2024), 32 pages. doi:10.1145/3678545

[35] Jiyang Li, Lin Huang, Siddharth Shah, Sean J Jones, Yincheng Jin, Dingran Wang, Adam Russell, Seokmin Choi, Yang Gao, Junsong Yuan, et al. 2023. Signring: Continuous american sign language recognition using imu rings and virtual imu data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–29.

[36] Yunzhi Li, Vimal Mollyn, Kuang Yuan, and Patrick Carrington. 2024. WheelPoser: Sparse-IMU Based Body Pose Estimation for Wheelchair Users. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 8, 17 pages. doi:10.1145/3663548.3675638

[37] Dawei Liang, Guihong Li, Rebecca Adaimi, Radu Marculescu, and Edison Thomaz. 2022. AudioIMU: Enhancing Inertial Sensing-Based Activity Recognition with Acoustic Models. In *Proceedings of the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(ISWC '22)*. Association for Computing Machinery, New York, NY, USA, 44–48. doi:10.1145/3544794.3558471

[38] Fengzhou Liang, Tian Min, Chengshuo Xia, and Yuta Sugiura. 2025. Vsens Toolkit: AR-based Open-ended System for Virtual Sensors. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 754, 5 pages. doi:10.1145/3706599.3721177

[39] Gabriel Lipkowitz and Joseph Desimone. 2024. Palette-PrintAR: augmented reality design and simulation for multicolor resin 3D printing. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 349, 12 pages. doi:10.1145/3613904.3642909

[40] Jiazhou Liu, Barrett Ens, Arnaud Prouzeau, Jim Smiley, Isobel Kara Nixon, Sarah Goodwin, and Tim Dwyer. 2023. DataDancing: An Exploration of the Design Space For Visualisation View Management for 3D Surfaces and Spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 379, 17 pages. doi:10.1145/3544548.3580827

[41] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. 2015. SMPL: a skinned multi-person linear model. *ACM Trans. Graph.* 34, 6, Article 248 (Oct. 2015), 16 pages. doi:10.1145/2816795.2818013

[42] Akhil Mathur, Tianlin Zhang, Sourav Bhattacharya, Petar Velickovic, Leonid Joffe, Nicholas D. Lane, Fahim Kawsar, and Pietro Lio. 2018. Using Deep Data Augmentation Training to Address Software and Hardware Heterogeneities in Wearable and Smartphone Sensing Devices. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. 200–211. doi:10.1109/IPSN.2018.00048

[43] Jianyuan Min and Jinxiang Chai. 2012. Motion graphs++: a compact generative model for semantic motion analysis and synthesis. *ACM Trans. Graph.* 31, 6, Article 153 (nov 2012), 12 pages. doi:10.1145/2366145.2366172

[44] Vimal Mollyn, Riku Arakawa, Mayank Goel, Chris Harrison, and Karan Ahuja. 2023. IMUPoser: Full-Body Pose Estimation using IMUs in Phones, Watches, and Earbuds. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 529, 12 pages. doi:10.1145/3544548.3581392

[45] Seungwhan Moon, Andrea Madotto, Zhaojiang Lin, Alireza Dirafzoon, Aparajita Saraf, Amy Bearman, and Babak Damavandi. 2022. IMU2CLIP: Multimodal Contrastive Learning for IMU Motion Sensors from Egocentric Videos and Text. arXiv:2210.14395 [cs.CV] https://arxiv.org/abs/2210.14395

[46] Muhammad Naeem, Wilson Ozuem, Kerry Howell, and Silvia Ranfagni. 2023. A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods* 22 (2023), 16094069231205789. doi:10.1177/16094069231205789 arXiv:https://doi.org/10.1177/16094069231205789

[47] Francisco Javier Ordóñez and Daniel Roggen. 2016. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* 16, 1 (2016). doi:10.3390/s16010115

[48] I. Orha and S. Oniga. 2014. Study regarding the optimal sensors placement on the body for human activity recognition. In *2014 IEEE 20th International Symposium for Design and Technology in Electronic Packaging (SIITME)*. 203–206. doi:10.1109/SIITME.2014.6967028

[49] Farshid Salemi Parizi, Eric Whitmire, and Shwetak Patel. 2020. AuraRing: Precise Electromagnetic Finger Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 4, Article 150 (Sept. 2020), 28 pages. doi:10.1145/3369831

[50] Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas, and Michael J. Black. 2019. Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. 10975–10985.

[51] Ling Pei, Songpengcheng Xia, Lei Chu, Fanyi Xiao, Qi Wu, Wenxian Yu, and Robert Qiu. 2021. MARS: Mixed Virtual and Real Wearable Sensors for Human Activity Recognition With Multidomain Deep Learning Model. *IEEE Internet of Things Journal* 8, 11 (2021), 9383–9396. doi:10.1109/JIOT.2021.3055859

[52] Patrick Reipschläger, Frederik Brudy, Raimund Dachselt, Justin Matejka, George Fitzmaurice, and Fraser Anderson. 2022. AvatAR: An Immersive Analysis Environment for Human Motion Data Combining Interactive 3D Avatars and Trajectories. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 23, 15 pages. doi:10.1145/3491102.3517676

[53] Attila Reiss. 2012. PAMAP2 Physical Activity Monitoring. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5NW2H.

[54] Téo Sanchez, Baptiste Caramiaux, Pierre Thiel, and Wendy E. Mackay. 2022. Deep Learning Uncertainty in Machine Teaching. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 173–190. doi:10.1145/3490099.3511117

[55] Panneer Selvam Santhalingam, Parth Pathak, Huzefa Rangwala, and Jana Kosecka. 2023. Synthetic smartwatch imu data generation from in-the-wild asl videos. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (2023), 1–34.

[56] Soyong Shin, Juyong Kim, Eni Halilaj, and Michael J. Black. 2024. WHAM: Reconstructing World-grounded Humans with Accurate 3D Motion. In *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*.

[57] Ken Shoemake. 1985. Animating rotation with quaternion curves. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '85)*. Association for Computing Machinery, New York, NY, USA, 245–254. doi:10.1145/325334.325242

[58] David Strömbäck, Sangxia Huang, and Valentin Radu. 2020. MM-Fit: Multimodal Deep Learning for Automatic Exercise Logging across Sensing Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 168 (dec 2020), 22 pages. doi:10.1145/3432701

[59] Yu Sun, Wu Liu, Qian Bao, Yili Fu, Tao Mei, and Michael J. Black. 2022. Putting People in their Place: Monocular Regression of 3D People in Depth. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 13233–13242. doi:10.1109/CVPR52688.2022.01289

[60] Megha Thukral, Sourish Gunesh Dhekane, Shruthi K. Hiremath, Harish Haresamudram, and Thomas Ploetz. 2025. Layout-Agnostic Human Activity Recognition in Smart Homes through Textual Descriptions Of Sensor Triggers (TDOST). *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 1, Article 19 (March 2025), 38 pages. doi:10.1145/3712278

[61] Catherine Tong, Jinchen Ge, and Nicholas D. Lane. 2022. Zero-Shot Learning for IMU-Based Activity Recognition Using Video Embeddings. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 180 (Dec. 2022), 23 pages. doi:10.1145/3494995

[62] Akihisa Tsukamoto, Naoto Yoshida, Tomoko Yonezawa, Kenji Mase, and Yu Enokibori. 2023. Where Are the Best Positions of IMU Sensors for HAR? - Approach by a Garment Device with Fine-Grained Grid IMUs -. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing* (Cancun, Quintana Roo, Mexico) *(UbiComp/ISWC '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 445–450. doi:10.1145/3594739.3610736

[63] Chiuan Wang, Hsuan-Ming Yeh, Bryan Wang, Te-Yen Wu, Hsin-Ruey Tsai, Rong-Hao Liang, Yi-Ping Hung, and Mike Y. Chen. 2016. CircuitStack: Supporting Rapid Prototyping and Evolution of Electronic Circuits. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 687–695. doi:10.1145/2984511.2984527

[64] Yihong Wu, Lingyun Yu, Jie Xu, Dazhen Deng, Jiachen Wang, Xiao Xie, Hui Zhang, and Yingcai Wu. 2023. AR-Enhanced Workouts: Exploring Visual Cues for At-Home Workout Videos in AR Environment. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 121, 15 pages. doi:10.1145/3586183.3606796

[65] Chengshuo Xia, Xinrui Fang, Riku Arakawa, and Yuta Sugiura. 2022. VoLearn: A Cross-Modal Operable Motion-Learning System Combined with Virtual Avatar and Auditory Feedback. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 81 (July 2022), 26 pages. doi:10.1145/3534576

[66] Chengshuo Xia, Ayane Saito, and Yuta Sugiura. 2022. Using the virtual data-driven measurement to support the prototyping of hand gesture recognition interface with distance sensor. *Sensors and Actuators A: Physical* 338 (2022), 113463. doi:10.1016/j.sna.2022.113463

[67] Chengshuo Xia and Yuta Sugiura. 2022. Virtual IMU Data Augmentation by Spring-Joint Model for Motion Exercises Recognition without Using Real Data. In *Proceedings of the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(ISWC '22)*. Association for Computing Machinery, New York, NY, USA, 79–83. doi:10.1145/3544794.3558460

[68] Fanyi Xiao, Ling Pei, Lei Chu, Danping Zou, Wenxian Yu, Yifan Zhu, and Tao Li. 2021. A Deep Learning Method for Complex Human Activity Recognition Using Virtual Wearable Sensors. In *Spatial Data and Intelligence*, Xiaofeng Meng, Xing Xie, Yang Yue, and Zhiming Ding (Eds.). Springer International Publishing, Cham, 261–270.

[69] Vasco Xu, Chenfeng Gao, Henry Hoffmann, and Karan Ahuja. 2024. MobilePoser: Real-Time Full-Body Pose Estimation and 3D Human Translation from IMUs in Mobile Consumer Devices. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 70, 11 pages. doi:10.1145/3654777.3676461

[70] Hui Ye and Hongbo Fu. 2022. ProGesAR: Mobile AR Prototyping for Proxemic and Gestural Interactions with Real-world IoT Enhanced Spaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 130, 14 pages. doi:10.1145/3491102.3517689

[71] A. D. Young, M. J. Ling, and D. K. Arvind. 2011. IMUSim: A simulation environment for inertial sensing algorithm design and evaluation. In *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks*. IEEE, 199–210.

[72] Shibo Zhang, Yuqi Zhao, Dzung Tri Nguyen, Runsheng Xu, Sougata Sen, Josiah Hester, and Nabil Alshurafa. 2020. NeckSense: A Multi-Sensor Necklace for Detecting Eating Activities in Free-Living Conditions. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 72 (June 2020), 26 pages. doi:10.1145/3397313

[73] Zhen Zhang, Samuel Abraham, Alex Lee, Yichen Li, Gota Morota, Dong Sam Ha, and Sook Shin. 2024. SegIt: Empowering Sensor Data Labeling with Enhanced Efficiency and Security. In *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI) (MLMI '24)*. Association for Computing Machinery, New York, NY, USA, 97–103. doi:10.1145/3696271.3696287

[74] Haozhe Zhou, Riku Arakawa, Yuvraj Agarwal, and Mayank Goel. 2025. IMUCoCo: Enabling Flexible On-Body IMU Placement for Human Pose Estimation and Activity Recognition. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. ACM, 1–16. doi:10.1145/3746059.3747695

[75] Zhongyi Zhou and Koji Yatani. 2021. Enhancing Model Assessment in Vision-based Interactive Machine Teaching through Real-time Saliency Map Visualization. In *Adjunct Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 112–114. doi:10.1145/3474349.3480194

[76] Zhongyi Zhou and Koji Yatani. 2022. Gesture-aware Interactive Machine Teaching with In-situ Object Annotations. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 27, 14 pages. doi:10.1145/3526113.3545648