On a Real Real-Time Wearable Human Activity Recognition System

Hui Liu[a], Tingting Xue[b] and Tanja Schultz[c]
Cognitive Systems Lab, University of Bremen, Germany

Keywords: Human Activity Recognition, Real-Time Systems, Wearables, Biodevices, Biosignals, Digital Signal Processing, Biosignal Processing, Machine Learning, Performance Analysis, Plug-and-Play.

Abstract: Many human activity recognition (HAR) systems have the ultimate application scenarios in real-time, while most literature has limited the HAR study to offline models. Some mentioned real-time or online applications, but the investigation of implementing and evaluating a real-time HAR system was missing. With our years of experience developing and demonstrating real-time HAR systems, we brief the implementation of offline HAR models, including hardware specifications, software engineering, data collection, biosignal processing, feature study, and human activity modeling, and then focus on the transition from offline to real-time models for details of window length, overlap ratio, sensor/device selection, feature selection, graphical user interface (GUI), and on-the-air functionality. We also indicate the evaluation of a real-time HAR system and put forward tips to improve the performance of wearable-based HAR.

1 BACKGROUND

Human activity recognition (HAR) is increasingly becoming a hot research topic and a technology that assists in all aspects of life. High-quality sensory observations applicable to recognizing users' activities and behaviors, including electrical, magnetic, mechanical (kinetic), optical, acoustic, thermal, and chemical biosignals, are inseparable from sensors' sophisticated design and appropriate application (Liu et al., 2023). Related research is emerging, which can be divided into two main categories through the application of sensing technologies — external sensing and internal sensing (Lara and Labrador, 2012). The latter is the object of this paper, which can provide users with unrestricted movement volume and wearable daily application experience.

HAR for medical and rehabilitation analysis, behavior and habit understanding, or activity modeling for game figures does not require real-time: a stable offline system with a high accuracy rate for processing stored biosignals is sufficient to meet the demand. In contrast, device control and human-machine interaction scenarios mostly call for the implementation of real-time HAR, such as game control, interactive user interface, sports assistance, and abnormal motion recognition, among others. The vast majority of wearable sensor-based HAR research papers explore offline models. Few mentions, outlooks, or preliminarily investigates the prospect of real-time applications. It is worth pointing out that some research publications use the name *real-time* loosely or inaccurately; Instead, their systems or approaches should be better referred to as *online* systems.

A strictly defined real-time recognition program must guarantee a response within specified time constraints, often referred to as "deadlines" (Ben-Ari, 2006). It controls an environment by receiving data, processing them, and returning the results sufficiently quickly to affect the environment at that time (Greenberger, 1965). In the perspective of digital signal processing (DSP) of real-time recognition systems, the analyzed (input) and generated (output) samples should be processed (or generated) continuously in the time it takes to input and output the same set of samples independent of the processing delay (Kuo et al., 2013). Consider a simple quantitative example; if an HAR system requires more than 1 second to process, recognize, and respond to a 1-second window/frame of recorded signals, the user will feel the system's recognition outputs slower and slower. Thus, after a short duration, it becomes a "congested" system that disrupts the real-time experience. If window overlapping, an often applied technology in HAR for improving recognition accuracy, is taken into ac-

[a] https://orcid.org/0000-0002-6850-9570
[b] https://orcid.org/0000-0002-5815-7217
[c] https://orcid.org/0000-0002-9809-7028

Liu, H., Xue, T. and Schultz, T.
On a Real Real-Time Wearable Human Activity Recognition System.
DOI: 10.5220/0011927700003414
In *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023) - Volume 5: HEALTHINF*, pages 711-720
ISBN: 978-989-758-631-6; ISSN: 2184-4305

count, the processing time for each input required for a real-time recognition system is even more demanding. Moreover, even though it is possible to provide more computation time by widening the window length, thus satisfying real-time in terms of the definition (without causing cumulative computation delay), the immediacy of HAR related to the interactive interface and device control does not tolerate a large window size. Just imagine that a gamer performs the "jump" activity to hit the question mark brick, and after two seconds, a crisp gold coin sounds.

We introduced our real-time HAR system with its on-the-fly functionality at the beginning of 2019, which the academic community highly recognized (Liu and Schultz, 2019). In the following four years, this system was invited to live demonstrations and interactive presentations on over 50 academic and industrial occasions, and was invariably well-received by the scholars and the engineers in attendance. From these dozens of hours of live real-time show-how, we have gradually validated the proposed approach's legitimacy, the implemented software interface's practicality, and the realized recognition system's real-time performance from the system running and the third-party feedback. We gathered suggestions, directions of interest, and other reflections from the scientific and industrial communities. Given this, we share and analyze all the technical details, design experiences, and gained perspectives of our real-time HAR system in this paper in order to provide reference to peer researchers and to be further validated by them.

2 OFFLINE MODELS

In the context of a research project in collaboration with industry, we integrated various wearable sensors in a medical knee bandage and used them for HAR experiments with the aim of providing a technological aid for post-operative rehabilitation and protection. Following the state-of-the-art HAR research pipeline (Liu et al., 2022a), we started with the equipment and setup study on the basis of the application scenarios.

2.1 Hardware Specifications: Devices, Settings, Carrier, and Wearable Sensor Integration

After testing different wearable devices, we applied *biosignalsplux Researcher Kit*[1] as the biosignal recording device that provides expandable solutions of hot-swappable sensors and automatic synchronization, since we need to place many kinds of sensors at different body positions around the knee in our HAR research under the framework of the collaborative project.

Several preliminary in-house HAR research efforts have validated the feasibility and stability of the biosignalsplux hubs and their attached selectable sensors, such as electromyography sensors (EMG), accelerometers (ACC), and electrogoniometers (EMG) in relevant research tasks from multiple perspectives (Rebelo et al., 2013) (Palyafári, 2015). In addition, we have included in our series of follow-up acquisition tasks (see Section 5) additional biomechanical and bioacoustic sensors, i.e., gyroscopes (GYRO) and microphones (MIC), which were produced specially for us by the provider. The gyroscope has been proven effective for HAR in the extensive literature (Ha and Choi, 2016) (Barna et al., 2019), while indications in the individual literature showed that the airborne microphone might be used to recognition workshop activity (Lukowicz et al., 2004).

Each *biosignalsplux* hub can simultaneously acquire up to 8 channels of signals from arbitrarily selected sensors, with a maximum sampling rate of 1000 Hz and a maximum quantization level of 16 bits. Three hub-channels are required for ACC and GYRO, respectively; EGM is dual-channel; each bipolar EMG, as well as the MIC, occupies one channel. For the three acquisition periods we have performed so far (Liu and Schultz, 2018) (Liu and Schultz, 2019) (Liu et al., 2021a), two hubs (up to 16 channels) or three hubs (up to 24 channels) were used, depending on the number of sensors employed. The hubs can be synchronized automatically via a signal synchronization cable. Due to the data volume limitation of Bluetooth real-time transmission, a sampling rate of 1000 Hz was adopted for four EMG sensors and one airborne microphone, while other biomechanical signals, ACCs, GYROs, and the EGM, were sampled at 100 Hz and upsampled to align EMG.

With the support of our research project partners, we applied *Bauerfeind*'s *Genutrain* model knee bandage[2] as a carrier for wearable sensors (see Figure 1). The detailed integration scheme is listed in Table 1.

2.2 Software Implementation and Data Corpora

In order to expediently control multiple wireless devices, stably acquire and archive multimodal biosignals, and smoothly perform up to 24-dimensional

[1] https://www.pluxbiosignals.com/products/researcher-kit (accessed January 18, 2023)

[2] https://www.bauerfeind.de/de/produkte/bandagen/knie/details/product/genutrain (accessed January 18, 2023)

Figure 1: The *Bauerfeind*'s *Genutrain* model knee bandage.

Table 1: Sensor integration scheme with positions and measured muscles.

Sensor	**Placement/Muscle**
ACC 1 (upper)	Thigh, proximal ventral
ACC 2 (lower)	Shank, distal ventral
GYRO 1 (upper)	Thigh, proximal ventral
GYRO 2 (lower)	Shank, distal ventral
EMG 1 (upper-front)	Musculus vastus medialis
EMG 2 (lower-front)	Musculus tibialis anterior
EMG 3 (upper-back)	Musculus biceps femoris
EMG 4 (lower-back)	Musculus gastrocnemius
EGM (lateral)	Knee of the right leg
MIC (lateral)	Knee of the right leg

real-time visualization in preparation for subsequent real-time recognition systems, we did not use the open-source acquisition software provided by the sensor company, as many studies do; Instead, we utilized the software development kit (SDK) and implemented our software, Activity Signal Kit (ASK).

The ASK baseline software can connect to wearable biosignal recording devices automatically, enable multi-sensor data acquisition and archiving, apply protocol-for-pushbutton mechanism of practical segmentation and annotation, process multichannel biosignals, extract features, and model human activities with the iteration of training-recognition-evaluation (Liu, 2021). A series of upgraded and expanded versions of the baseline software, such as the plug-and-play version (see Section 3.2), android mobile version, and virtual reality version, have been developed on the foundation of the robust baseline software with real-time HAR functionality (see Section 3.1).

By applying the developed software, three datasets of human movements were gradually collected and made public. The pilot dataset CSL17 (1 subject, 7 activities of daily living, 15 minutes) was used for validating the software implementation and the HAR research workflow (Liu and Schultz, 2018), based on which the current stable real-time HAR system runs stably. The advanced dataset CSL18 (4 subjects, 18 activities of daily living and sports, 90 minutes) and the comprehensive dataset CSL-SHARE (20 subjects, 22 types of activities of daily living and sports, 691 minutes) were utilized successfully for further offline HAR model research and will be applied for the future person-independent real-time HAR system (Liu and Schultz, 2019) (Liu et al., 2021a).

The activities involved in these datasets include standing, sitting, standing up, sitting down, one-leg jumping, two-leg jumping, walking, walking in a curve (left/right), walking upstairs, walking downstairs, left/right facing, lateral shuffling (left/right), jogging, and V-cutting (left/right).

A human activity dataset of this size can no longer be handled by trivial means and call for data mining, machine learning, and data analysis (Weiner et al., 2017) (see Section 2.3).

2.3 Biosignal Processing, Feature-Related Research and Activity Modeling

Several digital signal processing (DSP) tasks occur in the early stages of research, even during acquisition, such as amplification, filtering, and denoising. Normalization can be performed on the whole collected biosignals; however, real-time systems need to use accumulated normalization, for only a continuous influx of short-term streams is available. We did not focus on the DSP approaches before windowing, for they are more hardware-based and device-related. We conducted a series of window-based (DSP) experiments, especially on sensor selection (Liu, 2021), feature stacking and feature space reduction (Hartmann et al., 2020) (Hartmann et al., 2021), feature selection (Liu, 2021), and high-level feature design (Hartmann et al., 2022a), to ensure the no-deep learning's minimum sensor group and efficient feature representations.

Besides, some basic parameters of DSP, such as window length and overlap ratio, also need to be optimized in the process of parameter tuning, for which we conducted an analysis of human activity duration and concluded that a typical single motion of an average human body is basically between 1—2 seconds and is normally distributed in the population (Liu and Schultz, 2022).

We used hidden Markov models (HMM) (Ames, 1989) to model human activities (Xue and Liu, 2022). On the pilot dataset, each activity was modeled with one HMM state, achieving high accuracy on the seven-class person-dependent recognition. Increasing the HMM for each activity to the same number of HMM states was subsequently studied. Related literature (Rebelo et al., 2013) using ten HMM states provided a reference. We investigated one to ten HMM states for person-independent recognition and concluded that using eight HMMs for each activity yielded the best results on the applied experimental data. A question arose: It may be a reasonable choice to describe walking with several states, but is it necessary to have eight states for standing or sitting down? We solved the problem by proposing *Motion Units* to model human activities (Liu et al., 2021b):

- Could/should each activity contain a separate, explanatory number of states?
- Is there an approach to design HMMs of human activities more rule-based, normalized over blindly "trying"?

Inspired by kinesiological knowledge and the concept of the phoneme in ASR, each activity is composed of a different number of *Motion Units*. *Motion Units* are shareable among relevant activities, and the whole human activity modeling scheme based on MUs is highly interpretable, generalizable, and expandable.

3 REAL-TIME SYSTEM

Real-time systems are not necessarily the ultimate goal of HAR. Meanwhile, a great offline HAR model does not certainly fit a real-time system. Therefore, applying HAR in real-time scenarios as a way forward requires many aspects that must be investigated.

3.1 From Offline Towards Online

Many hardware, software, model, and parameter adjustments play pivotal roles in moving from offline to real-time HAR recognition.

3.1.1 Window Length and Overlap Ratio

The window length and the Ratio length not only affect the offline HAR modeling but are also two of the most critical parameters for the real-time HAR performance in an actual system, which impact at least the following aspects that have relationships between each other: recognition accuracy, response speed, computational cost, and (real) real-time. By linking window length and the overlap ratio, a shorter step size results in a shorter delay of the recognition outcomes, but the interim recognition results may fluctuate due to temporary search errors. On the other hand, longer delay due to long step sizes contradicts the characteristics of a real-time system, though it generates more accurate interim recognition results (Liu and Schultz, 2019).

In (Liu and Schultz, 2022), we concluded through statistical analysis that a normal, healthy single motion lasts about 1—2 seconds and is normally distributed among the population, which is the a priori information benefiting the parameter-tuning experiments of window length and overlap ratio to optimize the balance of recognition accuracy versus processing delay.

The results of our iterative experiments and statistical analysis can provide the following experience as a reference for a real-time HAR system that recognizes single motions (e.g., walking forward/upstairs/downstairs/in-curve, standing up, sitting down, one-leg/two-leg jumping, and squatting, among others) and stable postures (e.g., standing, sitting, and lying, among others): a window length of 400 ms and an overlap ratio of 50% (i.e., 200 ms).

Moreover, the choice of window length and overlap ratio is also related to other factors, such as the features chosen, the task of recognition (immediate interaction versus accurate analysis), and the terminal/server performance, among others. The two parameters must be experimented with comprehensively for each specific research orientation and application scenario.

3.1.2 Sensor and Device Selection

Offline sensor selection experiments help identify effective sensors for HAR and uncover those that are redundant or of little help (Liu, 2021). However, real-time systems are geared toward real-world applications, so further narrowing down the types and numbers of sensors is a more beneficial process. In addition, the choice of simpler, smaller, and lighter devices contributes to the ease of use of the technology. A practical experience: Experiments show that applying both ACCs and GYROs on the upper and lower leg improves recognition accuracy, but taking away GYRO results in only a tiny decrease in accuracy (Liu, 2021). Therefore, combined with the requirements of small-scale and wearability, a completely wireless *MuscleBAN*[3] that does not contain GYRO

[3]https://www.pluxbiosignals.com/products/muscleban-kit (accessed December 20, 2022)

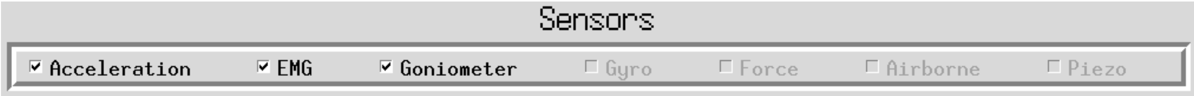

Figure 2: The “Sensors” option in the main menu of our real-time HAR system.

Figure 3: The “Activities” option in the main menu of our real-time HAR system.

and has ACC and EMG sensors can be a better choice for a (real) real-time HAR system.

In our real-time HAR system, various sensors are free to be connected and selected to test the performance of different sensors and their combinations in the actual recognition process (see Figure 2).

3.1.3 Feature Selection

Deep learning uses deep feature representations. However, recent literature suggests that deep learning-based recognition does not work principally superior to the application of handcrafted features (Bento et al., 2022), which has been validated on various datasets. For example, on the CSL-SHARE dataset (Liu et al., 2021a), the recognition accuracy achieved by the *Motion Units*-based HMM modeling (Liu et al., 2021b) mentioned in Section 2.3 outperformed the deep neural network algorithm and its variations (Mekruksavanich et al., 2022). The same happens as well for other datasets (Hu et al., 2018) (Micucci et al., 2017). Equally important, *Motion Units* activity modeling is entirely interpretable, and it is straightforward to model new activities, just like adding new words in automatic speech recognition, which makes training and recognition efficient.

Non-deep learning involves the procedure of feature extraction. The time consumption of the feature computation significantly impacts the real-time recognition, yielding a time complexity investigation. We applied the Time Series Feature Extraction Library (TSFEL) (Barandas et al., 2020), which has been proven effective and efficient in previous work on multimodal biosignal processing (Naseeb and Saeedi, 2020) and our previous signal processing studies (Liu et al., 2022b) (Liu et al., 2022c). In (Rodrigues et al., 2022) and (Liu, 2021), the authors listed and applied most features with low computational complexity in the temporal, statistical, and frequency domains from TSFEL. Such kind of study provided valuable references for real-time HAR systems.

3.1.4 Graphical User Interface Design

We developed a graphical user interface (GUI) for real-time HAR recognition on the ASK baseline software. In addition to the “sensors‘” option, as mentioned in Section 3.1.2, the activities to be recognized can also be selected in the main menu (see Figure 3).

In the real-time recognition window, the left side visualizes the real-time multichannel biosignal curves of all installed and selected sensors, and the right side continuously displays the window-based recognition results (see Figure 4). Our in-house HMM recognizer BioKIT (Telaar et al., 2014) automatically provides the *n*-Best recognition results, where typically, *n* is set to 3, i.e., three possible recognition results are illustrated. A prominent color indicates the activity with the highest probability.

A well-run plotting animation for real-time HAR is visually stimulating and appealing to the viewer. However, visualization is not always indispensable. For example, when interacting or controlling with HAR, people tend to focus solely on whether the recognition results can be used to accurately interoperate with the system and do not care whether the resulting biosignals are visualized or the different results are described in detail.

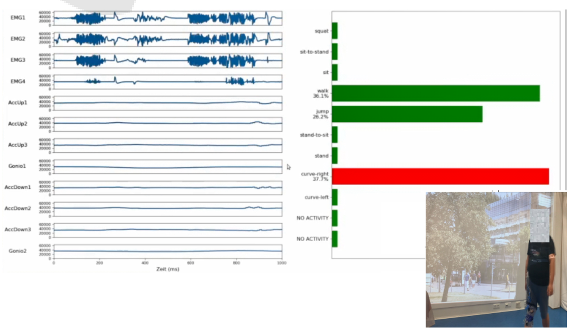

Figure 4: Screenshot of the real-time HAR interface. Left: the visualization of multimodal biosignal acquisition; right: the 3-best recognition results; bottom right: a video recording of the corresponding activity. The video was synthesized and was not a software feature.

3.2 Novel Online Functionality: On-the-Air Plug-and-Play

After achieving a stable performance of the basic real-time HAR that initially contains seven daily activities (sitting, standing, sitting down, standing up, walking forward, walking a left-turn curve, walking a right-turn curve), we have innovated a plug-and-play function, which is designed to easily and quickly add new recognizable activities to the system or provide new data to existing activities. The user can input the (new or existing) name of the activity in the main menu and enter the “Annotate” mode (see Figure 5). The system will start multimodal biosignal collection and the ”protocol-for-pushbutton” mechanism for automatic data segmentation and annotation. After about one minute of recording, the real-time HAR system is restarted, and the newly collected data are trained together with the initial corpus: either an existing activity’s model is updated with additional data, or a new activity is ready for recognition.

In dozens of live demonstrations over the past four years, this on-the-fly add-on function has always been well received. During the interactive sessions, scholars and engineers from different domains proposed various new activities to challenge the system’s robustness. It is worth mentioning that none of the new activities proposed by the audience failed to be recognized.

All proposed activities from third parties are listed and counted in Figure 6, with the purpose of helping researchers know which everyday whole-body activities, in the opinion of people in different fields, should be valuable for a real-time HAR system but rarely exist in HAR datasets.

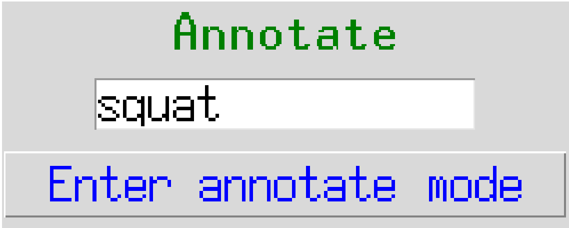

Figure 5: The “Annotate” option in the main menu of our real-time HAR system.

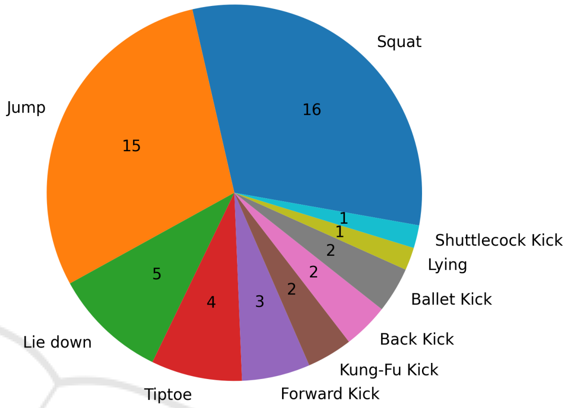

Figure 6: Proposed activities from third parties for challenging the plug-and-play add-on in our real-time HAR system during different demonstration events. The numbers in the pie chart deliver how many times each activity was proposed and tested.

4 EVALUATION OF REAL-TIME PERFORMANCE

Recent research shows that validation methods can influence HAR mobile systems (Bragança et al., 2022). Like machine learning algorithms’ evaluation in other fields, offline HAR models can be evaluated and validated through quantitative metrics and processes, including:

- Different types of recognition rates, such as macro average accuracy.
- Different types of error rates.
- Precision values.
- Recall values.
- F1-scores.
- Confusion matrices.
- Different types of cross-validation approaches, such as n-fold cross-validation and leave-one-out cross-validation (LOOCV).

However, the abovementioned metrics and processes cannot be directly grafted into evaluating real-time HAR systems.

4.1 Pseudo-Real-Time Evaluation

The evaluation of a real-time (recognition) system and real-time evaluations (RTE) are two different things. A straightforward evaluation scheme for real-time HAR can be realized in a pseudo-real-time way:

- If the training dataset is a field collection without strict protocol and annotated well, any data piece of it can be used to simulate a real-time data acquisition to evaluate the system’s performance in real-time recognition by the metrics listed above. Limitations: In this case, the pseudo-real-time validation is similar or even the same as the offline model evaluation, providing few research values.
- The training dataset, in most cases, is recorded following rigorous or relatively strict acquisition

protocols. For example, in one session of one participant, one particular activity or activity sequence is specified to be acquired several times orderly. Manual synthesis is required to simulate a pseudo-real-time time series with such a dataset, such as extracting various slices of the same participant in different recorded sessions, splicing them, and reannotating the generated pseudo data. Limitations: the synthesis cannot guarantee that the artificial data can highly simulate the data sequences recorded in a realistic real-time system.

Pseudo-real time, though practical to implement, is not a substitute for real real-time evaluation.

4.2 Is the System Real Real-Time?

The observations and subjective perceptions of the experimenter and audience are not negligible real-time evaluation criteria. A system that is consistently quickly responsive but inaccurate, or a system that is accurate but generates increasing recognition delays, can be perceived by the naked eye in a certain period on the basis of intuitive and detailed result visualization. In addition, quantitative analysis can be done objectively to facilitate the researcher's judgment. We recorded the time applied for each recognition completion in the backend of our system, and the average response time was 84.72 ± 10.09 ms per recognition (20 sessions, each over 120 minutes). Since the window length and the overlap ratio used in our system for optimal performance are 400 ms and 50%, respectively, the step size of the real-time recognition is 200 ms. According to the comparison between 84.72 ms and 200 ms, our system does not output recognition results more and more slowly because of the passage of time, which reflects its real-time capability from one perspective.

It is essential to point out that, unlike the analysis metrics of the offline model, the real-time performance analysis metrics are machine-dependent. The configuration of the host machine, the energy mode (battery/power supply), the system setups, the biosignal devices' battery level, the wireless (Bluetooth/WLAN) transmission stability, and the visualization's resource consumption, among others, all affect the response time of the real-time HAR system. We used a 6-year-old *Intel Core i7* laptop with average configurations, an external power supply, integrated graphics, 16 GB RAM, and Bluetooth 3.0. Therefore, the values measured on this machine objectively exhibit that the real-time performance of our HAR system should be pervasive to a great extent.

4.3 Is the System Robust?

We answer this question objectively with a series of quantitative metrics, aiming to provide some ideas and references.

Our system has been demonstrated more than **50 times** so far in public events such as academic conferences, project meetings, industrial exhibitions, and science fairs. The events range from a small talk of less than an hour to a technology booth of several hours. Based on the most conservative estimate, the system has been publicly demonstrated for more than **1,500 minutes** (30 minutes $\times$ 50 times). This duration does not include pre-event preparation rehearsals, regular system maintenance, and other tests. At all times, the system has never failed to work in any way.

We performed **20 sessions** of continuous recognition runs, each lasting over **120 minutes**. That means a total of over **1,440 minutes**. The efficiency and accuracy of the recognition were consistently validated and ensured during these trials (see Chapter 4.2). The time of 120 minutes has a substantial reference value, considering the upper limit of the battery supply of the wearable biosignal acquisition equipment.

Our plug-and-play add-on has also been challenged more than **50 times** so far, with a total of **10 new activities** proposed by third-party observers, each of which has been perfectly recognized. Plug-and-play multimodal data acquisition and automatic segmentation and annotation mechanism have always performed satisfactorily.

Admittedly, as with all software engineering tasks, large-scale and long-duration field-testing is the ultimate way to verify and assure the quality of a real-time HAR system. The above subsections aim to supply researchers with some experience for the evaluation of the designed real-time HAR system prior to being able to conduct field-testing.

5 TIPS FOR ENHANCING REAL-TIME PERFORMANCE

Based on our years of experience, we offer some hands-on means to improve the performance of real-time wearable systems. They play crucial roles not only in research or demonstration, but also in facilitating biosignal acquisition. After all, data collection is always in real-time.

- Connect mobile host devices, such as laptops, tablets, and cell phones, to a continuous external power supply instead of batteries.
- Configure optimal performance in the operating

system instead of the standard or power-saving mode.

- Charge the batteries of wearable biosignal acquisition devices fully and continuously observe their battery volume.
- Minimize interference on Bluetooth or WLAN transmissions, including signal interference and strong magnetic field interference (e.g., acquisition devices too close to a charger).
- Pay attention to the distance of wireless transmissions.
- Turn off the irrelevant communication ports on the host machine, such as switching off the wireless network connection when Bluetooth is used for acquiring data.
- Set a reasonable step size for visualization. Plotting animation is resource-consuming.
- Care about the sensor situations constantly, such as electrodes for bioelectrical sensing. Common problems include electrode detachment/switch, no/problematic grounding, and dry electrodes.

The above experiences are referential for real-time data acquisition, experiment, and demonstration. Once a real-time HAR system is put into practical application, it is impossible to control real users' behavior.

6 CONCLUSION AND OUTLOOK

A large portion of HAR systems has the ultimate application scenarios in real-time. Unfortunately, most of the literature has limited the study of HAR to offline models. Some mentioned real-time applications, but not necessarily to practice, hone and validate real real-time qualities. With our years of experience developing and demonstrating real-time HAR systems, we first introduced the implementation of offline HAR models and then focused on the transition from offline to real-time models. The evaluation of a real-time HAR system and tips to improve the performance of wearable-based HAR are also contributions of this paper.

Based on the introduced ASK baseline software and its plug-and-play, ASK 2.0 software focusing on interactive real-time machine learning is under development for open-source sharing, and a preview version has already yielded good results (Hartmann et al., 2022b). Meanwhile, segment-based instead of window-based real-time HAR is also a direction worth investigating, for which automatic segmentation by subsequence search (Folgado et al., 2022) or change point detection (Rodrigues et al., 2022) by self-similarity matrix can be used as novel input sources of real-time training and recognition to improve the accuracy. Last but not least, applying real-time HAR systems to a broader stage, such as fall detection and human-machine interaction in the metaverse, is a future research prospect that many scientists strive for.

ACKNOWLEDGMENTS

The research reported in this paper has been partially supported by the German Federal Ministry of Education and Research; Project-ID 16DHBKI047 "IntEL4CoRo - Integrated Learning Environment for Cognitive Robotics", University of Bremen.

REFERENCES

Ames, C. (1989). The markov process as a compositional model: A survey and tutorial. *Leonardo*, 22(2):175–187.

Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T., and Gamboa, H. (2020). TSFEL: Time series feature extraction library. *SoftwareX*, 11:100456.

Barna, A., Masum, A. K. M., Hossain, M. E., Bahadur, E. H., and Alam, M. S. (2019). A study on human activity recognition using gyroscope, accelerometer, temperature and humidity data. In *2019 international conference on electrical, computer and communication engineering (ecce)*, pages 1–6. IEEE.

Ben-Ari, M. (2006). *Principles of concurrent and distributed programming*. Pearson Education.

Bento, N., Rebelo, J., Barandas, M., Carreiro, A. V., Campagner, A., Cabitza, F., and Gamboa, H. (2022). Comparing handcrafted features and deep neural representations for domain generalization in human activity recognition. *Sensors*, 22(19):7324.

Bragança, H., Colonna, J. G., Oliveira, H. A., and Souto, E. (2022). How validation methodology influences human activity recognition mobile systems. *Sensors*, 22(6):2360.

Folgado, D., Fernandes, Barandas, M., Antunes, M., Nunes, M. L., Liu, H., Hartmann, Y., Schultz, T., and Gamboa, H. (2022). TSSEARCH: Time series subsequence search library. *SoftwareX*, 18:101049.

Greenberger, M. (1965). Programming real-time computer systems.

Ha, S. and Choi, S. (2016). Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In *2016 international joint conference on neural networks (IJCNN)*, pages 381–388. IEEE.

Hartmann, Y., Liu, H., Lahrberg, S., and Schultz, T. (2022a). Interpretable high-level features for human activity recognition. In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 4: BIOSIGNALS*, pages 40–49.

Hartmann, Y., Liu, H., and Schultz, T. (2020). Feature space reduction for multimodal human activity recognition. In *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 4: BIOSIGNALS*, pages 135–140. INSTICC, SciTePress.

Hartmann, Y., Liu, H., and Schultz, T. (2021). Feature space reduction for human activity recognition based on multi-channel biosignals. In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 215–222. INSTICC, SciTePress.

Hartmann, Y., Liu, H., and Schultz, T. (2022b). Interactive and interpretable online human activity recognition. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 109–111. IEEE.

Hu, B., Rouse, E., and Hargrove, L. (2018). Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals. *Frontiers in Robotics and AI*, 5:14.

Kuo, S. M., Lee, B. H., and Tian, W. (2013). *Real-time digital signal processing: fundamentals, implementations and applications*. John Wiley & Sons.

Lara, O. D. and Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209.

Liu, H. (2021). *Biosignal Processing and Activity Modeling for Multimodal Human Activity Recognition*. PhD thesis, Universität Bremen.

Liu, H., Gamboa, H., and Schultz, T. (2023). Sensor-based human activity and behavior research: Where advanced sensing and recognition technologies meet. *Sensors*, 23(1):125.

Liu, H., Hartmann, Y., and Schultz, T. (2021a). CSL-SHARE: A multimodal wearable sensor-based human activity dataset. *Frontiers in Computer Science*, 3:90.

Liu, H., Hartmann, Y., and Schultz, T. (2021b). Motion Units: Generalized sequence modeling of human activities for sensor-based activity recognition. In *EUSIPCO 2021 — 29th European Signal Processing Conference*, pages 1506–1510. IEEE.

Liu, H., Hartmann, Y., and Schultz, T. (2022a). A practical wearable sensor-based human activity recognition research pipeline. In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5: HEALTHINF*, pages 847–856.

Liu, H., Jiang, K., Gamboa, H., Xue, T., and Schultz, T. (2022b). Bell shape embodying zhongyong: The pitch histogram of traditional chinese anhemitonic pentatonic folk songs. *Applied Sciences*, 12(16):8343.

Liu, H. and Schultz, T. (2018). ASK: A framework for data acquisition and activity recognition. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 3: BIOSIGNALS*, pages 262–268. INSTICC, SciTePress.

Liu, H. and Schultz, T. (2019). A wearable real-time human activity recognition system using biosensors integrated into a knee bandage. In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 1: BIODEVICES*, pages 47–55. INSTICC, SciTePress.

Liu, H. and Schultz, T. (2022). How long are various types of daily activities? statistical analysis of a multimodal wearable sensor-based human activity dataset. In *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5: HEALTHINF*, pages 680–688.

Liu, H., Xue, T., and Schultz, T. (2022c). Merged pitch histograms and pitch-duration histograms. In *Proceedings of the 19th International Conference on Signal Processing and Multimedia Applications - SIGMAP*, pages 32–39. INSTICC, SciTePress.

Lukowicz, P., Ward, J. A., Junker, H., Stäger, M., Tröster, G., Atrash, A., and Starner, T. (2004). Recognizing workshop activity using body worn microphones and accelerometers. In *In Pervasive Computing*, pages 18–32.

Mekruksavanich, S., Jantawong, P., and Jitpattanakul, A. (2022). A deep learning-based model for human activity recognition using biosensors embedded into a smart knee bandage. *Procedia Computer Science*, 214:621–627.

Micucci, D., Mobilio, M., and Napoletano, P. (2017). UniMiB SHAR: A dataset for human activity recognition using acceleration data from smartphones. *Applied Sciences*, 7(10):1101.

Naseeb, C. and Saeedi, B. A. (2020). Activity recognition for locomotion and transportation dataset using deep learning. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, pages 329–334.

Palyafári, R. (2015). Continuous activity recognition for an intelligent knee orthosis; an out-of-lab study. Master's thesis, Karlsruher Institut für Technologie.

Rebelo, D., Amma, C., Gamboa, H., and Schultz, T. (2013). Human activity recognition for an intelligent knee orthosis. In *BIOSIGNALS 2013 - 6th International Conference on Bio-inspired Systems and Signal Processing*, pages 368–371.

Rodrigues, J., Liu, H., Folgado, D., Belo, D., Schultz, T., and Gamboa, H. (2022). Feature-based information retrieval of multimodal biosignals with a self-similarity matrix: Focus on automatic segmentation. *Biosensors*, 12(12):1182.

Telaar, D., Wand, M., Gehrig, D., Putze, F., Amma, C., Heger, D., Vu, N. T., Erhardt, M., Schlippe, T., Janke, M., et al. (2014). Biokit—real-time decoder for biosignal processing. In *Fifteenth Annual Conference of the International Speech Communication Association*.

Weiner, J., Diener, L., Stelter, S., Externest, E., Kühl, S., Herff, C., Putze, F., Schulze, T., Salous, M., Liu, H., et al. (2017). Bremen big data challenge 2017: Predicting university cafeteria load. In *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, pages 380–386. Springer.

Xue, T. and Liu, H. (2022). Hidden markov model and its application in human activity recognition and fall detection: A review. In *Communications, Signal Processing, and Systems*, pages 863–869. Springer Singapore.