Effect of Sliding Window Sizes on Sensor-Based Human Activity Recognition Using Smartwatch Sensors and Deep Learning Approaches

Sakorn Mekruksavanich[1], Ponnipa Jantawong[1] and Anuchit Jitpattanakul[2,*]
[1]*Department of Computer Engineering, School of Information and Communication Technology*
University of Phayao, Phayao, Thailand
sakorn.me@up.ac.th and ponnipa.jantawong@gmail.com
[2]*Intelligent and Nonlinear Dynamic Innovations Research Center, Department of Mathematics*
Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
anuchit.j@sci.kmutnb.ac.th

2024 5th International Conference on Big Data Analytics and Practices (IBDAP) | 979-8-3503-9174-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/IBDAP62940.2024.10689691

***Abstract*—Smartwatch sensors for human activity recognition (HAR) have gained significant attention due to their applications in healthcare and fitness monitoring. The effectiveness of HAR systems largely depends on the choice of sliding window widths for sensor data segmentation. This study investigates the impact of varying sliding window widths on the accuracy of HAR using wristwatch sensors and deep learning techniques. We conducted experiments using the daily human activity (DHA) dataset, comprising sensor data from 11 distinct activities. Data was preprocessed and segmented using window sizes ranging from 5 to 40 seconds. Four deep learning models (CNN, LSTM, BiLSTM, and CNN-LSTM) were employed and evaluated using accuracy, precision, recall, and F1-score. Window size significantly affected HAR performance. Smaller windows improved short-duration activity recognition but increased computational complexity, while larger windows reduced computational load but decreased accuracy for rapid activity changes. The CNN-LSTM hybrid model consistently outperformed other models, achieving 92.11% accuracy with a 20-second window and overlapping segmentation. This research provides valuable insights into balancing recognition accuracy and computational resources in smartwatch sensor-based HAR, contributing to the development of efficient and accurate systems for real-world applications.**

***Keywords*—human activity recognition, smartwatch sensors, sliding window, deep learning, sensor data segmentation**

I. Introduction

Human activity recognition (HAR) has gained considerable prominence in the past few decades owing to its many uses for medical care [1]–[4], sports monitoring [5]–[8], and individual support [9]–[12]. Due to the progress in wearable devices, smartwatches have gained popularity for gathering sensor data about people's movements. These gadgets are equipped with various sensors, including accelerometers, gyroscopes, and magnetometers, that collect detailed data on the individual's motions and actions [13].

Deep learning algorithms have demonstrated exceptional effectiveness in HAR challenges, surpassing conventional machine learning techniques [14]–[16]. These algorithms, including convolutional neural networks (CNNs) and long-short-term memory (LSTM) networks, can acquire structured characteristics from unprocessed sensor data without requiring any human feature engineering [17]. A vital component of HAR utilizing deep learning involves pre-processing data collected by sensors. Dividing the uninterrupted flow of sensor data into windows of specific proportions is a widely used method to prepare the data for training and inference of models. The sliding window size selection may substantially affect the identification quality and the computing complexity of HAR applications.

Prior research has investigated the impact of different sliding window widths on HAR utilizing diverse sensor paradigms and deep learning algorithms [18], [19]. The study results indicate that employing large sliding windows is optional for achieving outstanding performance measures. This is because doing so would increase the cost of processing and the reaction capabilities of any application developed through the trained model [17]. Nevertheless, a more thorough examination is needed regarding the influence of varying window widths on the utilization of smartwatch detectors and cutting-edge deep learning methods.

This study examines how varying sliding window widths impact the accuracy of sensor-based HAR utilizing wristwatch sensors and deep learning methods. The information we gather is the daily human activity (DHA) dataset. It consists of sensor data collected from individuals engaged in 11 distinct activities, including strolling, jogging, seating, and standing. We operate many algorithms for deep learning, such as CNN, LSTM, bidirectional LSTM (BiLSTM), and CNN-LSTM, to categorize the actions based on the retrieved characteristics. We assess the efficacy of each approach by using criteria including accuracy, precision, recall, and F1-score.

This paper's primary contributions are outlined below:

1) Our study examines how varying sliding window widths impact the accuracy of HAR utilizing wristwatch sensors and advanced deep learning methods.

2) Our findings provide valuable insights into the practical implications of the trade-off between identification accuracy and computing complexity when determining sliding window widths for sensor-based HAR.

II. Related Works

The process of data segmentation in HAR, particularly the widely used sliding window approach, is a significant area of research 20], [21. This approach, known for its simplicity and reliability, has been the focus of numerous studies [22]. Researchers have explored a variety of window lengths in their investigations, underscoring the importance of this method in HAR.

Movements such as strolling, running, and moving up or down the stairs have been recognized using small window sizes, 0.5 s and 0.8 s [23]. The classification of stationary, walking, running, and biking movement phases has been performed using a decision tree classifier in combination with a window size of 1 second [24]. Furthermore, a time interval of 2 seconds, combined with a neural network [25], has been used to categorize different types of motion, such as strolling, going up and down stairs, running, and seated with different body positions. This approach has yielded a mean accuracy of 93%. Portable smartphones have been utilized for categorizing walking, standing, and ascending stairs with high accuracy scores of 84% utilizing larger window widths, such as 5 seconds, and numerous approaches [26]. The recognition of strolling, immobile, running, and cycling actions is achieved by employing a window size of 7.5 seconds when the smartphone is put in the individual's trouser pocket [27]. The approach, along with the K-nearest neighbor machine learning technique, results in a classification accuracy of 93.9%.

Investigators in deep learning for HAR have examined the effects of segmenting input data into window sizes as a pre-processing technique. Mairittha et al. [28] conducted data annotation for a motion detection technique employing inertial (acceleration and angular velocity) portable sensing in both simple-LSTM and hybrid CNN-LSTM models. The experimenters used a window size of 5.12 seconds (at a frequency of 20 Hz), resulting in about 100 frames with overlapping. Ebner et al. [29] introduced a new method that utilizes analytical conversions and artificially created sensor channels to recognize activities, namely acceleration and rotational velocity. The researchers conducted experiments with window widths of 2, 2.5, and 3 seconds at a sampling rate of 50 Hz. This corresponds to 100, 125, and 150 frames, respectively. The experiments were conducted without any overlapping. The authors' conclusion indicated that the impact of window size on accuracy was minimal, with a modest inclination towards decreased accuracy as window widths increased.

Despite the considerable research on the selection of window widths for HAR, there is a clear need for comprehensive studies that specifically focus on the effect of sliding window sizes when using smartwatch sensors and advanced deep-learning models. This study is designed to meet this need by examining the impact of varying sliding window widths on the accuracy of sensor-based HAR utilizing wristwatch sensors and deep learning methodologies.

III. Methodology

The sensor-based HAR architecture implemented in this study consists of four primary steps: data gathering, data pre-processing, data production, and model training and assessment, as seen in Fig. 1.

A. Daily Human Activity (DHA) Dataset

The DHA dataset [30] from Kookmin University contains smartwatch accelerometer data from two individuals performing 11 distinct activities over four weeks. Data was collected using an Apple Watch Series-2 at 10Hz and transmitted to an iPhone 6. Activities were conducted in office spaces (5), kitchens (3), and outdoors (3), with no concurrent activities assumed. The dataset includes information on the number of samples for each activity. The quantity of samples for each task is condensed in Table I.

TABLE I: A list of activities in DHA dataset

Activity	Abbrv.	Location	No. of data[a]
Office work	Ow	Office	62,711
Reading	Re	Office	36,976
Writing	Wr	Office	27,677
Taking a rest	Tr	Office	31,265
Playing a game	Pg	Office	51,906
Eating	Ea	Kitchen	46,155
Cooking	Co	Kitchen	10,563
Washing dishes	Wd	Kitchen	10,712
Walking	Wa	Outdoors	25,768
Running	Ru	Outdoors	6,452
Taking a transport	Tt	Outdoors	28,483

[a]Number of raw accelerometer data.

Fig. 2 shows 2D graphs of accelerometer data for 11 activities in the DHA dataset, displaying x, y, and z acceleration over time. The patterns vary distinctly across activities:

- Office work, reading, writing, resting, gaming, and eating show modest acceleration amplitudes.
- Walking and running display higher, more regular accelerations.
- Cooking, dish cleaning, and transportation exhibit mild acceleration fluctuations.

B. Data Pre-processing

The raw sensor data was subjected to noise reduction and normalization during the pre-processing step. The pre-processed sensor data was segmented by a fixed-width sliding window technique through completing these procedures.

This research used two segmentation techniques to examine the influence of different sliding window widths. The initial approach, known as the overlapping temporal window (OW), entails the application of a window of a predetermined size to the input data sequence. This window creates training and test samples, employing a particular validation method. Nevertheless, this approach results in substantial bias due

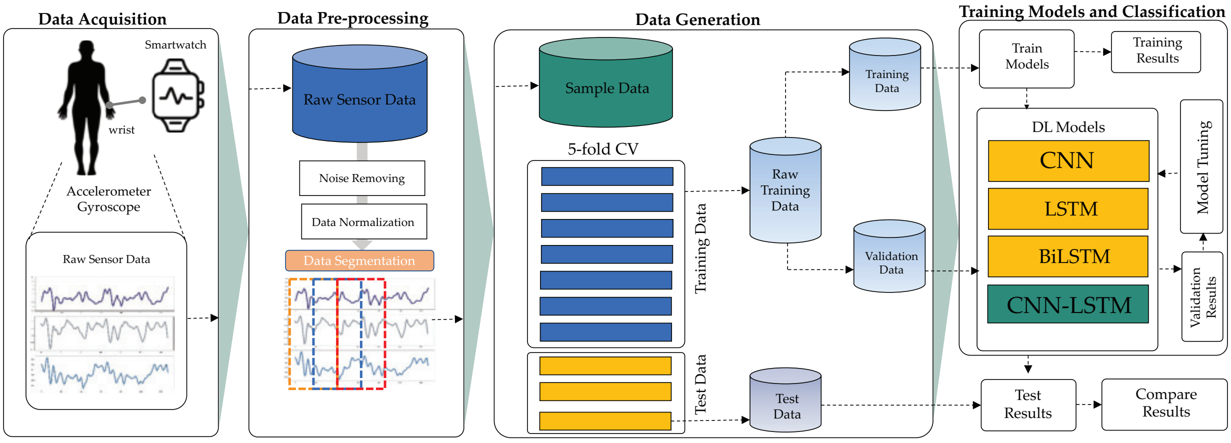


Fig. 1: The HAR framework based on smartwatch sensors used in this work.

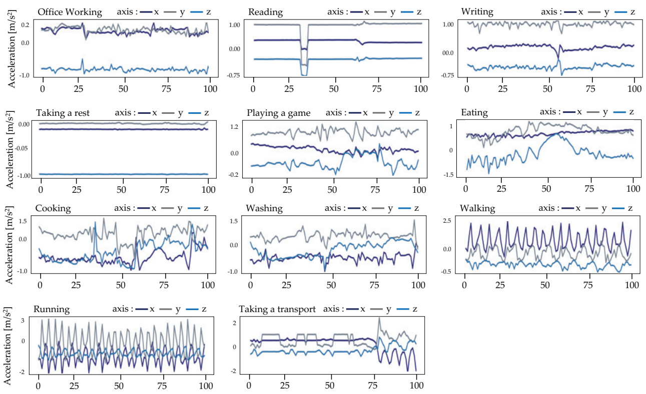


Fig. 2: Some samples of 11 daily human activities from DHA dataset.

to the 50% overlap between subsequent sliding windows. In order to reduce this bias, a different method called the non-overlapping temporal window (NOW) was used. The NOW strategy, as opposed to the OW method, has the drawback of producing fewer samples due to the lack of overlapping temporal frames. Fig. 3 illustrates two sample generation algorithms for segmenting sensor data, where X, Y, and Z correspond to the three portions of a tri-axial IMU sensor.

C. *Deep Learning Models*

This study examines the impact of different sliding window widths on sensor-based HAR employing smartwatch sensors and Deep Learning methods. This study utilizes four deep learning approaches: CNN, LSTM, BiLSTM, and CNN-LSTM. The structure of each model employed in this study is illustrated in Fig. 4.

IV. Experiments and Results

Our research investigates how varying sliding window sizes affect deep learning models' interpretation in identifying human movements from smartwatch sensor data. We utilized the DHA dataset to conduct our experiments. Four distinct neural network architectures were employed: CNN, LSTM, BiLSTM, and a hybrid CNN-LSTM model.

The experimentation process involved training and evaluating these models with different time windows. We explored durations spanning from 5 to 40 seconds. In our analysis, we applied both non-overlapping and overlapping segmentation approaches to the data. We relied on widely-used performance metrics to gauge the efficacy of each model configuration. These included accuracy, precision, recall, and F1-score. This comprehensive evaluation allowed us to assess the impact of window size on the models' ability to classify human activities based on wearable sensor input accurately.

A. *Experimental Findings*

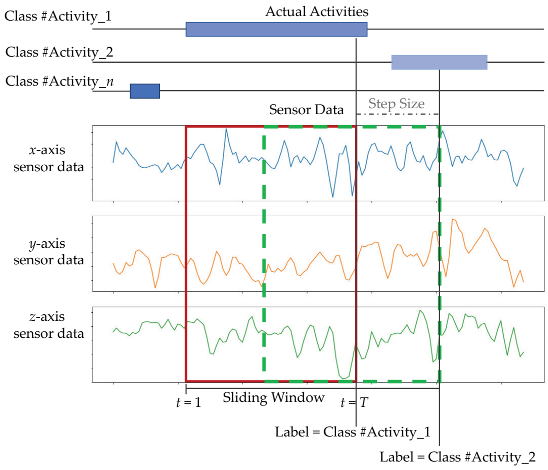
This research examines how altering the window size (5, 10, 20, 30, and 40 seconds) affects training diverse deep learning issues throughout various scenarios. Furthermore, we acknowledge that the recognition ability improves as the window size rises, especially for intricate actions, as seen in Table II and Table III.

Table II shows performance metrics for CNN, LSTM, BiLSTM, and CNN-LSTM models trained on non-overlapping data segments of 5 to 40 seconds. Key findings:

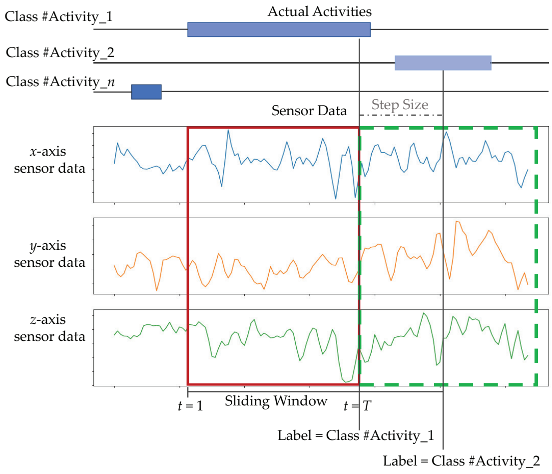
- CNN-LSTM outperforms other models across all window sizes.
- LSTM and BiLSTM performance declines with larger windows (20+ seconds).
- CNN maintains stable performance across window sizes.
- CNN-LSTM achieves highest accuracy (88.09%) at 5 seconds, but remains robust (¿86%) for larger windows.
- Precision, recall, and F1-score trends mirror accuracy.

Results emphasize the importance of window size and architecture selection in sensor-based activity recognition, with CNN-LSTM emerging as the most effective and robust approach.

Table III shows performance metrics for CNN, LSTM, BiLSTM, and CNN-LSTM models trained on 50% overlapping data segments of 5 to 40 seconds. Key findings:



(a) Overlapping temporal window (OW) with 50% overlap



(b) Non-overlapping temporal window (NOW)

Fig. 3: Fixed-width sliding window segmentation strategies: (a) overlapping temporal window (OW) with 50% overlap, and (b) non-overlapping temporal window (NOW) – in both strategies, X, Y, and Z denote the three components of the tri-axial accelerometer sensor data.

TABLE II: Performance metrics of the deep learning models (CNN, LSTM, BiLSTM, and CNN-LSTM) trained and tested with different window sizes (5, 10, 20, 30, and 40 seconds) using the non-overlapping temporal window (NOW) segmentation protocol

Window Sizes	Model	Accuracy	Precision	Recall	F1-score
5 s	CNN	86.37%	82.94%	83.52%	83.12%
	LSTM	85.90%	85.56%	83.59%	84.50%
	BiLSTM	86.59%	86.22%	83.26%	84.64%
	CNN-LSTM	88.09%	85.40%	86.41%	85.83%
10 s	CNN	85.53%	82.20%	85.43%	83.55%
	LSTM	81.31%	81.01%	74.32%	77.21%
	BiLSTM	86.50%	86.46%	84.71%	85.40%
	CNN-LSTM	86.50%	84.41%	85.93%	85.07%
20 s	CNN	82.57%	82.12%	84.17%	82.89%
	LSTM	73.30%	72.51%	66.45%	68.23%
	BiLSTM	81.16%	81.16%	77.76%	78.95%
	CNN-LSTM	87.18%	88.56%	87.46%	87.82%
30 s	CNN	83.24%	85.47%	90.17%	87.44%
	LSTM	64.01%	72.68%	52.72%	59.29%
	BiLSTM	71.20%	74.89%	67.64%	69.98%
	CNN-LSTM	86.08%	87.79%	91.42%	89.24%
40 s	CNN	83.21%	82.87%	87.02%	84.65%
	LSTM	62.65%	75.25%	65.66%	68.91%
	BiLSTM	69.97%	68.88%	72.03%	68.68%
	CNN-LSTM	86.41%	87.19%	90.71%	88.56%

TABLE III: Performance metrics of the deep learning models (CNN, LSTM, BiLSTM, and CNN-LSTM) trained and tested with different window sizes (5, 10, 20, 30, and 40 seconds) using the overlapping temporal window (OW) segmentation protocol with 50% overlapping proportion

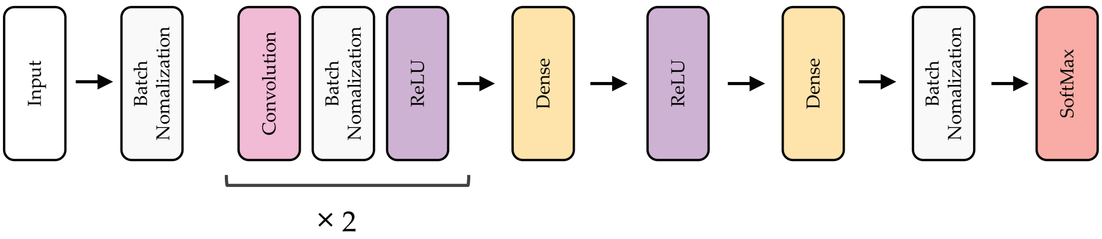
Window Sizes	Model	Accuracy	Precision	Recall	F1-score
5 s	CNN	89.21%	87.26%	88.19%	87.66%
	LSTM	88.79%	86.43%	86.89%	86.59%
	BiLSTM	90.98%	89.00%	90.30%	89.57%
	CNN-LSTM	91.55%	88.21%	90.85%	89.44%
10 s	CNN	87.94%	86.61%	85.41%	85.86%
	LSTM	87.40%	87.99%	85.64%	86.68%
	BiLSTM	88.87%	87.74%	86.48%	87.05%
	CNN-LSTM	91.91%	89.25%	90.93%	90.06%
20 s	CNN	86.38%	86.30%	89.49%	87.77%
	LSTM	79.23%	77.50%	75.13%	75.86%
	BiLSTM	84.43%	85.04%	81.14%	82.83%
	CNN-LSTM	92.11%	91.40%	90.49%	90.89%
30 s	CNN	86.48%	86.29%	89.18%	87.58%
	LSTM	70.39%	76.31%	65.26%	69.79%
	BiLSTM	77.93%	77.96%	72.54%	74.53%
	CNN-LSTM	90.91%	93.08%	93.57%	93.21%
40 s	CNN	85.76%	83.47%	91.61%	87.24%
	LSTM	59.36%	66.10%	49.27%	54.37%
	BiLSTM	71.69%	75.67%	70.80%	72.63%
	CNN-LSTM	91.31%	93.25%	95.44%	94.28%

- CNN-LSTM consistently outperforms other models across all window sizes.
- Overlapping protocol improves performance for all models compared to non-overlapping.
- LSTM performance declines sharply with more oversized windows (30+ seconds).
- BiLSTM and CNN maintain relatively stable performance across window sizes.
- CNN-LSTM achieves highest accuracy (92.11%) at 20 seconds.

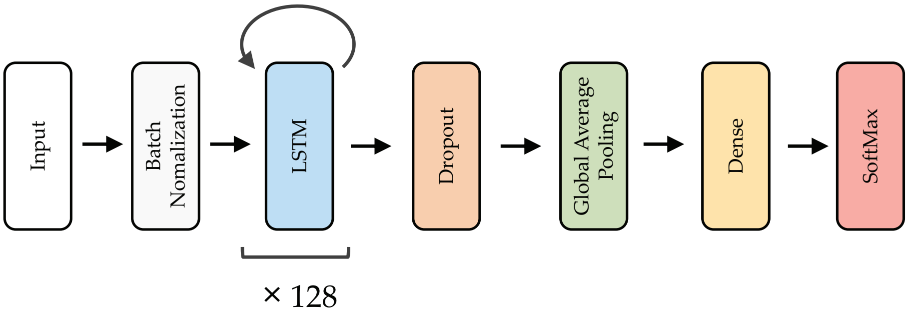
Results confirm CNN-LSTM's superiority for sensor-based activity recognition, especially with overlapping windows. Smaller window sizes generally lead to better performance. The overlapping protocol improves results by capturing more contextual information. Window size selection should consider sensor data characteristics and task requirements.

V. Conclusion and Future Works

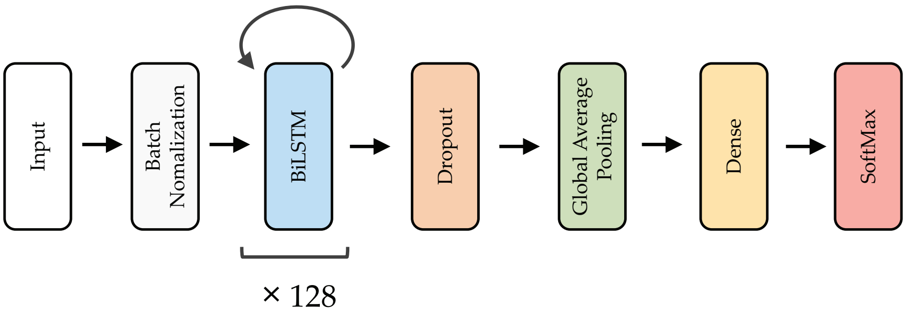
This study investigates how sliding window sizes affect deep learning models' performance in recognizing human activities



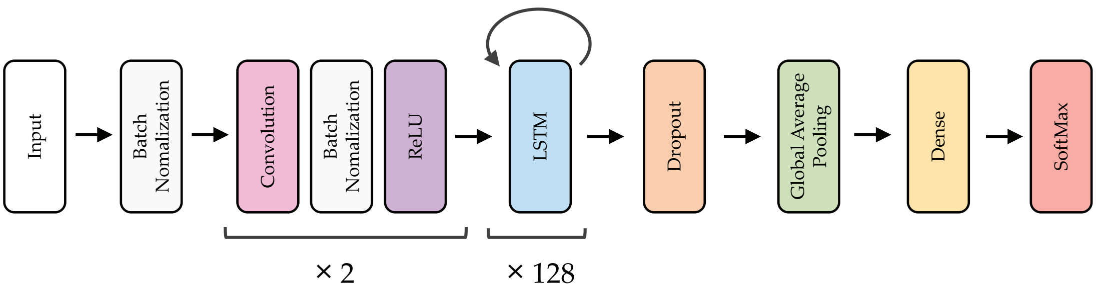
(a) CNN architecture



(b) LSTM architecture



(c) BiLSTM architecture



(d) CNN-LSTM architecture

Fig. 4: Deep learning architectures used in this study.

from smartwatch sensor data. Four models (CNN, LSTM, BiLSTM, CNN-LSTM) were tested on the DHA dataset using 5–40 second windows with overlapping and non-overlapping segmentation. Key findings:

- Smaller windows generally improved model performance.
- CNN-LSTM consistently outperformed other models.
- LSTM and BiLSTM were more sensitive to larger window sizes than CNN.
- Overlapping segmentation enhanced overall performance.

The CNN-LSTM hybrid model proved most effective and robust. Window size selection should consider sensor data characteristics, activity complexity, and application requirements.

Future work could explore different sensor modalities and fusion techniques, advanced architectures for complex dependencies, data augmentation for improved robustness, more extensive, more diverse datasets, and personalized activity recognition models.

Acknowledgment

This research project was supported by University of Phayao (Grant no. FF67-UoE-214); Thailand Science Research and Innovation Fund (Fundamental Fund 2024); National Science, Research and Innovation Fund (NSRF); and King Mongkut's University of Technology North Bangkok with Contract no. KMUTNB-FF-67-B-10.

References

[1] F. Demrozi, C. Turetta, P. H. Kindt, F. Chiarani, R. A. Bacchin, N. Valè, F. Pascucci, P. Cesari, N. Smania, S. Tamburin, and G. Pravadelli, "A low-cost wireless body area network for human activity recognition in healthy life and medical applications," *IEEE Transactions on Emerging Topics in Computing*, vol. 11, no. 4, pp. 839–850, 2023.

[2] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "Improving lower limb activity recognition based on inertial sensors using cnn-lstm network," in *2023 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C)*, 2023, pp. 170–173.

[3] P. Khan, Y. Kumar, and S. Kumar, "Capslstm-based human activity recognition for smart healthcare with scarce labeled data," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 1, pp. 707–716, 2024.

[4] S. Mekruksavanich and A. Jitpattanakul, "Deep residual neural network for aggressive physical activity recognition using surface electromyography sensors," in *2023 IEEE 14th International Conference on Software Engineering and Service Science (ICSESS)*, 2023, pp. 175–178.

[5] J.-S. Kim, "Dnn-based human activity recognition by learning initial motion data for virtual multi-sports," in *2021 23rd International Conference on Advanced Communication Technology (ICACT)*, 2021, pp. 373–375.

[6] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "Recognizing and understanding sport activities based on wearable sensor signals using deep residual network," in *2023 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C)*, 2023, pp. 166–169.

[7] J. Wei, B. Yu, H. Zhang, and J. Liu, "Skeleton based graph convolutional network method for action recognition in sports: A review," in *2023 8th IEEE International Conference on Network Intelligence and Digital Content (IC-NIDC)*, 2023, pp. 66–70.

[8] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "A hybrid deep learning neural network for recognizing exercise activity using inertial sensor and motion capture system," in *2023 4th International Conference on Big Data Analytics and Practices (IBDAP)*, 2023, pp. 1–5.

[9] S. Mekruksavanich and A. Jitpattanakul, "A lightweight deep residual network for recognizing activities in daily living using channel state information," in *2023 IEEE 14th International Conference on Software Engineering and Service Science (ICSESS)*, 2023, pp. 171–174.

[10] S. B. Rekha and M. V. Rao, "Methodical activity recognition and monitoring of a person through smart phone and wireless sensors," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, 2017, pp. 1456–1459.

[11] S. Kalimuthu, T. Perumal, R. Yaakob, E. Marlisah, and L. Babangida, "Human activity recognition based on smart home environment and their applications, challenges," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2021, pp. 815–819.

[12] S. Mekruksavanich and A. Jitpattanakul, "Classifying activities of electrical line workers based on deep learning approaches using wrist-worn sensor," in *2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2023, pp. 270–274.

[13] N. Hnoohom, A. Jitpattanakul, P. Inluergsri, P. Wongbudsri, and W. Ployput, "Multi-sensor-based fall detection and activity daily living classification by using ensemble learning," in *2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON)*, 2018, pp. 111–115.

[14] W. Sanpote, P. Jantawong, N. Hnoohom, A. Jitpattanakul, and S. Mekruksavanich, "Deep learning approaches for recognizing daily human activities using smart home sensors," in *2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, 2023, pp. 469–473.

[15] S. Abbas, G. A. Sampedro, S. Alsubai, S. Ojo, A. S. Almadhor, A. A. Hejaili, and L. Strazovska, "Advancing healthcare and elderly activity recognition: Active machine and deep learning for fine- grained heterogeneity activity recognition," *IEEE Access*, vol. 12, pp. 44 949–44 959, 2024.

[16] S. Mekruksavanich, P. Jantawong, and A. Jitpattanakul, "Deep learning approaches for har of daily living activities using imu sensors in smart glasses," in *2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, 2023, pp. 474–478.

[17] D. Noh, H. Yoon, and D. Lee, "A decade of progress in human motion recognition: A comprehensive survey from 2010 to 2020," *IEEE Access*, vol. 12, pp. 5684–5707, 2024.

[18] L. Sun, X. Yang, and C. Hu, "Dswhar: A dynamic sliding window based human activity recognition method," in *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*, 2022, pp. 1421–1426.

[19] M. H. M. Noor, Z. Salcic, and K. I.-K. Wang, "Dynamic sliding window method for physical activity recognition using a single tri-axial accelerometer," in *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)*, 2015, pp. 102–107.

[20] J. Wan, M. J. O'Grady, and G. M. P. O'Hare, "Dynamic sensor event segmentation for real-time activity recognition in a smart home context," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 287–301, Feb 2015.

[21] G. Wang, Q. Li, L. Wang, W. Wang, M. Wu, and T. Liu, "Impact of sliding window length in indoor human motion modes and pose pattern recognition based on smartphone sensors," *Sensors*, vol. 18, no. 6, 2018.

[22] S. Mekruksavanich and A. Jitpattanakul, "Rnn-based deep learning for physical activity recognition using smartwatch sensors: A case study of simple and complex activity recognition," *Mathematical Biosciences and Engineering*, vol. 19, no. 6, pp. 5671–5698, 2022.

[23] J.-H. Wang, J.-J. Ding, Y. Chen, and H.-H. Chen, "Real time accelerometer-based gait recognition using adaptive windowed wavelet transforms," in *2012 IEEE Asia Pacific Conference on Circuits and Systems*, 2012, pp. 591–594.

[24] L. Sun, D. Zhang, B. Li, B. Guo, and S. Li, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations," in *Ubiquitous Intelligence and Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 548–562.

[25] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in *2010 5th International Conference on Future Information Technology*, 2010, pp. 1–6.

[26] Y.-S. Lee and S.-B. Cho, "Activity recognition using hierarchical hidden markov models on a smartphone with 3d accelerometer," in *Hybrid Artificial Intelligent Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 460–467.

[27] P. Siirtola and J. Röning, "User-independent human activity recognition using a mobile phone: Offline recognition vs. real-time on device recognition," in *Distributed Computing and Artificial Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 617–627.

[28] N. Mairittha, T. Mairittha, and S. Inoue, "On-device deep personalization for robust activity data collection," *Sensors*, vol. 21, no. 1, 2021.

[29] M. Ebner, T. Fetzer, M. Bullmann, F. Deinzer, and M. Grzegorzek, "Recognition of typical locomotion activities based on the sensor data of a smartphone in pocket or hand," *Sensors*, vol. 20, no. 22, 2020.

[30] M.-C. Kwon and S. Choi, "Recognition of daily human activity using an artificial neural network and smartwatch," *Wireless Communications and Mobile Computing*, vol. 2018, no. 1, p. 2618045, 2018.