TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

Deep Learning Networks for Complex Activity Recognition Based on Wrist-worn Sensor

Sakorn Mekruksavanich[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
[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

***Abstract*—Wearable smart devices, such as smartphones and smartwatches, offer great potential as platforms for automated human action identification. However, accurately monitoring complex human actions on these devices poses a challenge due to the presence of similarities in patterns across different actions. This occurs when distinct human actions exhibit comparable signal patterns or characteristics. The placement of motion sensors on the body plays a crucial role in detecting human behavior. Typically, wearable sensors placed at the trouser pocket or a similar location are used for this purpose. However, this positioning is not suitable for identifying actions involving manual gestures. To address this, wrist-worn motion sensors are employed to detect these specific behaviors. This study aims to investigate the effectiveness of deep learning models in accurately categorizing complex human actions using sensor data from wrist-worn devices. Nine deep learning models utilizing convolutional neural networks and recurrent neural networks were examined for their identification capabilities. The models were evaluated using the WHARF dataset, a publicly available benchmark dataset for human activity recognition. The investigation revealed that the proposed CNN-BiGRU model outperformed other deep learning models, achieving an accuracy rate of 87.20% and an F1-score of 84.46%.**

***Keywords*—complex activity recognition, human activity recognition, wrist-worn sensor, deep learning, hybrid deep learning**

I. Introduction

The advancements in wearable technologies have made smartwatches a valuable tool for universal computing, particularly in managing wellness in daily life. Smartwatches, worn on the wrist, are electronic devices equipped with multiple sensors that enable continuous monitoring of an individual's activities [1]. This technology has found beneficial applications in various areas, such as medical care tracking [2] , fitness and sports monitoring [3], and behavior management [4]. In the realm of medical care, the ability to identify physical activities through sensor data from smartwatches has played a significant role in mitigating the negative effects of unhealthy lifestyles. For instance, monitoring an individual's eating-related activities has been recognized as a potentially valuable factor in treating various illnesses, including cancer, diabetes, obesity, and heart attacks [5], [6].

Individual bodily activity refers to the various states of the human body, including but not limited to running, strolling, and resting. Human activity recognition (HAR) is a current area of scholarly investigation that centers on the automated identification or evaluation of a specific individual user's actions through the analysis of relevant sensor data [7], [8]. The process of identifying activities through wearable sensors placed on different body areas or embedded in personal devices such as smartphones, sports bracelets, and smartwatches is known as sensor-based activity recognition [9], [10].

The use of wearable sensors in HAR has traditionally posed challenges in classifying multivariate time-series data. Feature extraction plays a crucial role in addressing this challenge and can be accomplished through statistical techniques in both the frequency and time domains [11]. Conventional machine learning techniques like Naïve Bayes, decision trees, and support vector machines have proven effective in accurately categorizing various human actions [12]. However, manual feature extraction requires specific domain knowledge or expertise, limiting their ability to detect distinctive characteristics for complex actions [13]. To overcome this limitation, deep learning approaches have employed convolutional neural networks (CNNs) to automatically extract abstract features from sensor data in the early stages of HAR research [14]. While CNNs excel in interpreting the spatial domain of sensor data and performing well for basic activities, they struggle to capture temporal characteristics essential for complex actions [15]. Consequently, recurrent neural networks (RNNs) are utilized in HAR to extract temporal information from wearable sensor data, addressing the temporal aspect [16]. Training RNNs poses a challenge due to vanishing or exploding gradients, which is addressed by the long short-term memory (LSTM) neural network architecture. Recent studies in HAR have successfully employed LSTM models to improve recognition abilities [17]. To mitigate the limitations of both CNNs and RNNs, hybrid deep learning models have been developed [18], [19].

This study presents a framework for recognizing complex human activities using sensor data collected from a wrist-worn wearable device. The framework aims to accurately identify intricate behaviors. The study evaluates five deep learning networks: CNN, LSTM, Bidirectional LSTM (BiLSTM), Gated Recurrent Unit (GRU), and Bidirectional GRU (BiGRU) in the context of complex activity recognition. Additionally, a hybrid deep learning model called CNN-BiGRU is proposed to improve identification effectiveness. The investigation utilizes the publicly available WHARF dataset, which includes accelerometer data capturing detailed activity information from sixteen individuals.

The subsequent sections of this paper are structured in the following manner. Section II provides an overview of recent literature on the topic. Section III provides a comprehensive account of the proposed model's specifics. The findings from the experiment are presented in Section IV. Section V concludes this study and presents potential areas for further investigation.

II. Related Works

In recent years, researchers have proposed various structures and techniques for identifying patterns in human movement. Many of these frameworks rely on wearable devices, particularly smartwatches, for activity recognition and tracking. This section provides important background information on studies related to HAR that utilize sensors, along with relevant research efforts in the field.

A. Complex Activity Recognition

In recent years, there has been a significant research focus on HAR using motion sensors embedded in wearable devices [20]. Additionally, there has been a growing interest in utilizing smartphone sensors for HAR in recent studies [12]. Numerous investigations have explored the identification of physical activities through wrist-worn devices. In a specific study [21], researchers investigated the role of sensors in smartwatches and smartphones for recognizing complex activities.

Trost et al. [22] utilized both a wrist-worn sensor and a hip sensor to identify seven specific physical activities. Logistic regression was employed as the classification algorithm. The authors successfully demonstrated the effectiveness of using wrist positioning for activity recognition. However, it's worth noting that the assessment of these two positions was carried out separately, without combining the data from both sensors. In another study [23], the authors utilized a single wrist-worn accelerometer to identify five different physical activities, including sitting, standing, and running. Additionally, in a related study [24], a wrist-worn accelerometer was used to identify a total of eight activities, including the complex task of computer usage .

In their study [25], the authors utilized a hidden Markov model along with a wrist-worn accelerometer and gyroscope to identify eating behavior. Eating activities were categorized into sub-activities such as resting, ingesting food, drinking liquids, using utensils, and related actions. The authors reported an accuracy of 84.30% for their approach. Similarly, Dong et al. [26] employed a combination of a wrist-worn accelerometer and gyroscope to identify eating behavior. They differentiated between eating and non-eating periods and achieved an accuracy level of 81%. Additionally, Sen et al. [27] used data from the accelerometer and gyroscope of a smartwatch to distinguish eating periods from comparable non-eating activities.

B. Deep Learning Approach for HAR

Despite the significant advancements achieved in HAR through traditional machine learning methods, certain constraints and disadvantages exist. Initially, the characteristics are obtained through a heuristic and manually designed approach heavily dependent on human or domain-specific expertise. Furthermore, it is worth noting that solely superficial characteristics can be derived based on human proficiency. The aforementioned superficial characteristics frequently pertain to various statistical parameters such as the average, variance, maximum, minimum, and so forth. The sensors cannot identify complex or contextually-relevant actions and are primarily suited for detecting basic movements such as strolling or jogging. Eating-related actions, such as the consumption of spaghetti or water, exhibit a high degree of complexity and prove challenging to discern by applying shallow features alone. The utilization of deep learning models has been observed to address the limitations mentioned above effectively [28]. Moreover, the deep neural network can extract high-level representations in deeper layers, rendering it more appropriate for intricate activity identification assignments.

The deep neural network (DNN) is an advanced form of the artificial neural network (ANN) that stands out for its increased depth. Unlike traditional ANNs, which have a limited number of hidden layers, DNNs are composed of a larger number of layers. This depth gives DNNs an advantage in learning as it allows them to handle extensive datasets more effectively. In complex models, DNNs are often employed as a condensed layer. For example, in CNNs, it is customary to incorporate multiple dense layers following the convolutional layers [29].

CNN is based on three fundamental concepts: sparse interactions, parameter sharing, and equivariant representations. Following the convolution process, subsequent layers typically include pooling and fully-connected layers that undertake regression or classification functions. CNN has demonstrated proficiency in extracting features from signals, yielding encouraging outcomes in image classification, speech recognition, and text investigation. CNN exhibits two benefits over alternative models when employed in the context of time series classification, such as in the case of the HAR task. These benefits are attributable to CNN's ability to capture local dependencies and scale invariance. The concept of local dependency in the context of HAR pertains to the likelihood of correlated signals nearby. On the other hand, scale invariance pertains to the property of remaining invariant across various rates or frequencies.

RNNs are widely used in speech recognition and processing of natural language to capture temporal relationships between neurons. One common practice is incorporating LSTM components into RNNs, where LSTM units serve as memory units through gradient descent. Although few studies have applied RNNs to HAR tasks [30], [31], these studies primarily aimed to improve learning efficiency and minimize resource usage.

The hybrid model combines multiple deep models, and a new hybrid approach that has gained attention recently is the fusion of CNNs and LSTM networks. In a referenced study, the combination of CNNs and LSTM models showed effective integration and improved performance compared to using only CNNs with dense layers [32]. Similar findings were reported in another studies [33]–[35]. CNNs excel at capturing spatial relationships, while RNNs can leverage temporal relationships. By integrating CNNs and RNNs, the model's capability to identify various activities with different durations and signal distributions can be enhanced.

III. The Proposed Framework

The present study employs a sensor-based HAR framework consisting of four primary stages: data acquisition, data pre-processing, data generation, and model training and assessment, as depicted in Fig. 1.

A. WHARF Dataset

The dataset utilized in this study comprises wrist-worn sensor data. It is publicly accessible under the name "ADL Recognition with Wrist-worn Accelerometer Dataset," commonly called the WHARF dataset [36]. The public can access the above resource through the UCI Machine Learning Repository [37]. The present dataset solely comprises accelerometer signals that were sampled at a frequency of 32 Hz. The dataset known as WHARF comprises 14 different activities, which are itemized in Table I. The previous actions were collected by a cohort of 16 volunteers, consisting of 11 male and 5 female individuals, whose ages ranged from 19 to 81 years. The individuals who participated were equipped with a triaxial accelerometer affixed to their right wrists, recording data at 32 Hz. The accelerometer readings are presented as a sequence of chronological data points, whereby the classification of actions is determined by the signal characteristics observed over a specific duration. Therefore, the present task pertains to the identification of time series data.

TABLE I
14 ACTIVITIES OF THE WHARF DATASET

Type	Activities
Bathroom use	Clean teeth
	Comb hair
Transferring	Rise from the bed
	Lie down on the bed
	Sit down on a chair
	Stand up from a chair
Feeding	Sip from a glass
	Dine with utensils
	Eat with spoon
	Pour water into a glass
Telephone usage capability	Make a telephone call
Mode of transportation	Climb the stairs
	Descend the stairs
	Walk

B. Data Pre-processing

The raw sensor data underwent several pre-processing steps. Initially, noise reduction was conducted by applying a median filter and a 3rd order low-pass Butterworth filter, with a cutoff frequency of 20 Hz. Furthermore, data normalization was performed using the Min-Max method. Afterward, the pre-processed sensor data were divided into segments using a sliding window approach. The window had a fixed width of 3 seconds, and an overlap ratio of 50% was applied.

C. The Proposed CNN-BiGRU Model

In this study, five different deep learning networks were investigated: CNN, LSTM, BiLSTM, GRU, and BiGRU. To improve recognition effectiveness, a new deep learning architecture called CNN-BiGRU was introduced. This architecture combines convolutional neural networks with bidirectional gated recurrent units. Fig. 2 provides a visual representation of the CNN-BiGRU design.

D. Training Deep Learning Models

In the context of deep learning, the model training procedure involves using hyperparameters to regulate the learning process. The present study employs a set of hyperparameters, namely: (i) the number of epochs, (ii) batch size, (iii) the rate of learning denoted by α, (iv) optimization, and (v) loss function, in the model that has been suggested. The hyperparameters were set by initializing the number of epochs to 200 and the batch size to 128. If no enhancements in the validation loss metric were observed following the completion of 30 epochs, we implemented an early stopping callback mechanism to terminate the training procedure. At the outset, the learning rate was established as α = 0.001. Subsequently, we increased it to 75% of its initial value if the accuracy of the validation of the suggested model did not exhibit any improvement over six consecutive epochs. In order to reduce the level of error, the Adam optimizer was employed with specific parameter values, namely, β_1 = 0.9, β_2 = 0.999, and $\epsilon = 1 \times 10^{-8}$. The categorical cross-entropy function is employed to compute the discrepancy in the optimizer.

E. Performance Measurement Criteria

The use of a confusion matrix is a valuable tool in assessing the recognition effectiveness of deep learning models, as it provides a simple and straightforward visualization of their effectiveness. The multi-class confusion matrix can be mathematically represented as a matrix where the rows correspond to the expected class instances and the columns correspond to the actual class instances.

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \cdots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \cdots & c_{nn} \end{bmatrix}$$

The confusion elements for each class are given by:
True positive: TP

$$TP(C_i) = C_{ii} \tag{1}$$

False positive: FP

$$FP(C_i) = \sum_{l=1}^{n} c_{li} - TP(C_i) \tag{2}$$

False negative: FN

$$FN(C_i) = \sum_{l=1}^{n} c_{il} - TP(C_i) \tag{3}$$

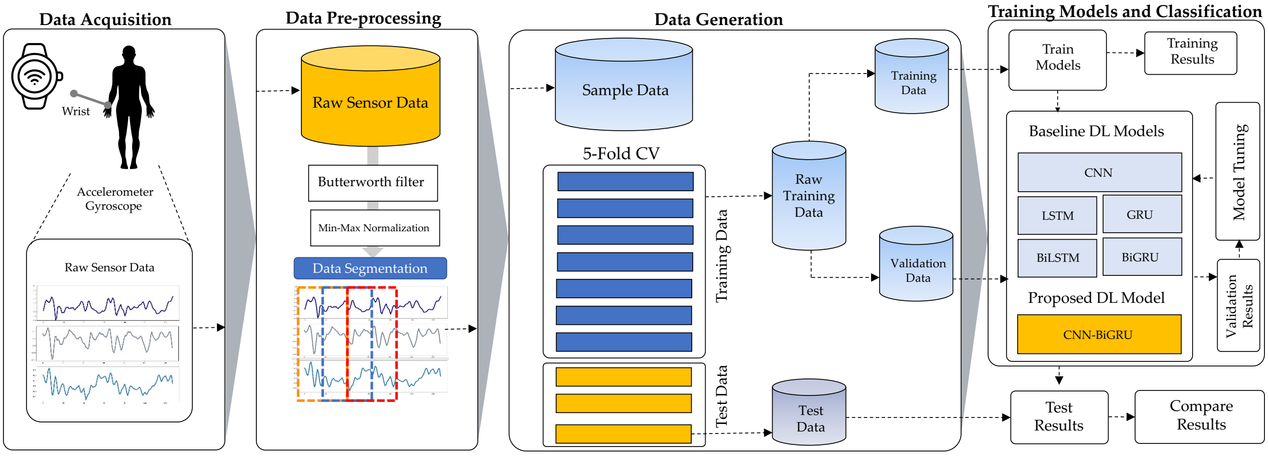

Fig. 1. The sensor-based framework based on wrist-worn sensors used in this work.

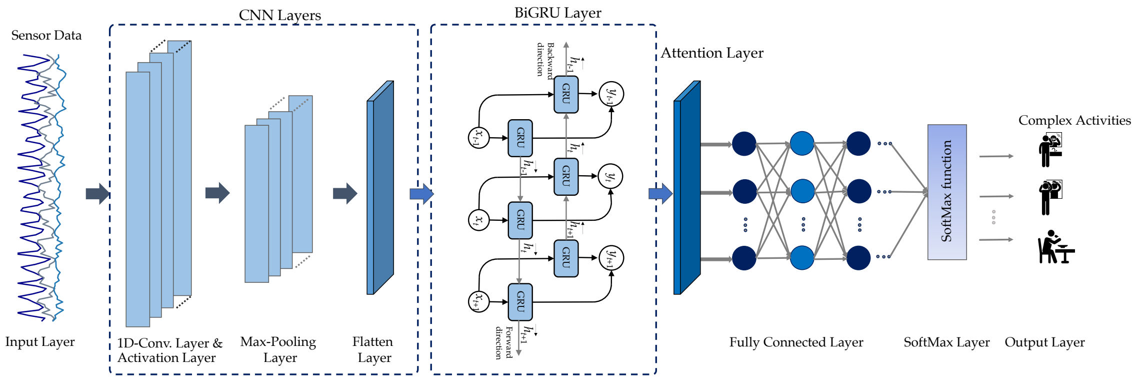

Fig. 2. The architecture of the proposed CNN-BiGRU.

True negative: TN

$$TN(C_i) = \sum_{l=1}^{n}\sum_{k=1}^{n} c_{lk} - TP(C_i) - FP(C_i) - FN(C_i) \quad (4)$$

The evaluation of the deep learning models employed in this research was conducted through a confusion matrix, which facilitated the computation of four conventional performance metrics, including accuracy. The evaluation of accuracy pertains to the depiction of systematic error. The computation involves determining the proportion of the combined number of true positives and true negatives to the overall quantity of entries.

Accuracy: Acc

$$Acc = \frac{1}{|Class|} \times \sum_{i=1}^{|Class|} \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i} \quad (5)$$

IV. Experiments

All experiments in this study were carried out on the Google Colab Pro+ platform using a Tesla V100. The development of the Python programming language involved the utilization of various libraries, including Python, TensorFlow, Keras, Scikit-Learn, Numpy, and Pandas. The purpose of this study is to investigate the identification capabilities of deep learning models based on CNN and LSTM architectures. Specifically, we aim to evaluate the impact of data sharing on the data segmentation process. Five different step sizes, namely 25, 50, 100, 150, and 200, were employed for this evaluation.

A. Experimental Findings

The WHARF dataset was created using a 5-fold cross-validation approach. In this study, we conducted a series of investigations to evaluate the performance of five key deep learning models: CNN, LSTM, BiLSTM, GRU, and BiGRU. Additionally, we introduced our proposed CNN-BiGRU model. Table II presents the recognition values achieved by each model, while Fig. 3 illustrates the confusion matrices for all models.

TABLE II
Identification Effectiveness of Deep Learning Models Using Sensor Data from the WHARF Dataset

Model	Recognition Performance		
	Accuracy%($\pm$SD%)	**Loss**	**F1-score%($\pm$SD%)**
CNN	82.35($\pm$0.53)	0.65($\pm$0.03)	78.33($\pm$0.99)
LSTM	74.95($\pm$1.63)	1.06($\pm$0.07)	70.55($\pm$1.88)
BiLSTM	71.91($\pm$1.19)	1.09($\pm$0.06)	67.37($\pm$1.64)
GRU	76.20($\pm$1.43)	0.73($\pm$0.03)	70.22($\pm$1.51)
BiGRU	78.66($\pm$0.96)	0.66($\pm$0.04)	73.61($\pm$0.99)
CNN-BiGRU	87.20($\pm$0.49)	0.50($\pm$0.04)	84.46($\pm$0.57)

The investigation revealed that the proposed CNN-BiGRU model achieved a peak accuracy of 87.20% and a peak F1-score of 84.46%. This model combines a CNN layer for

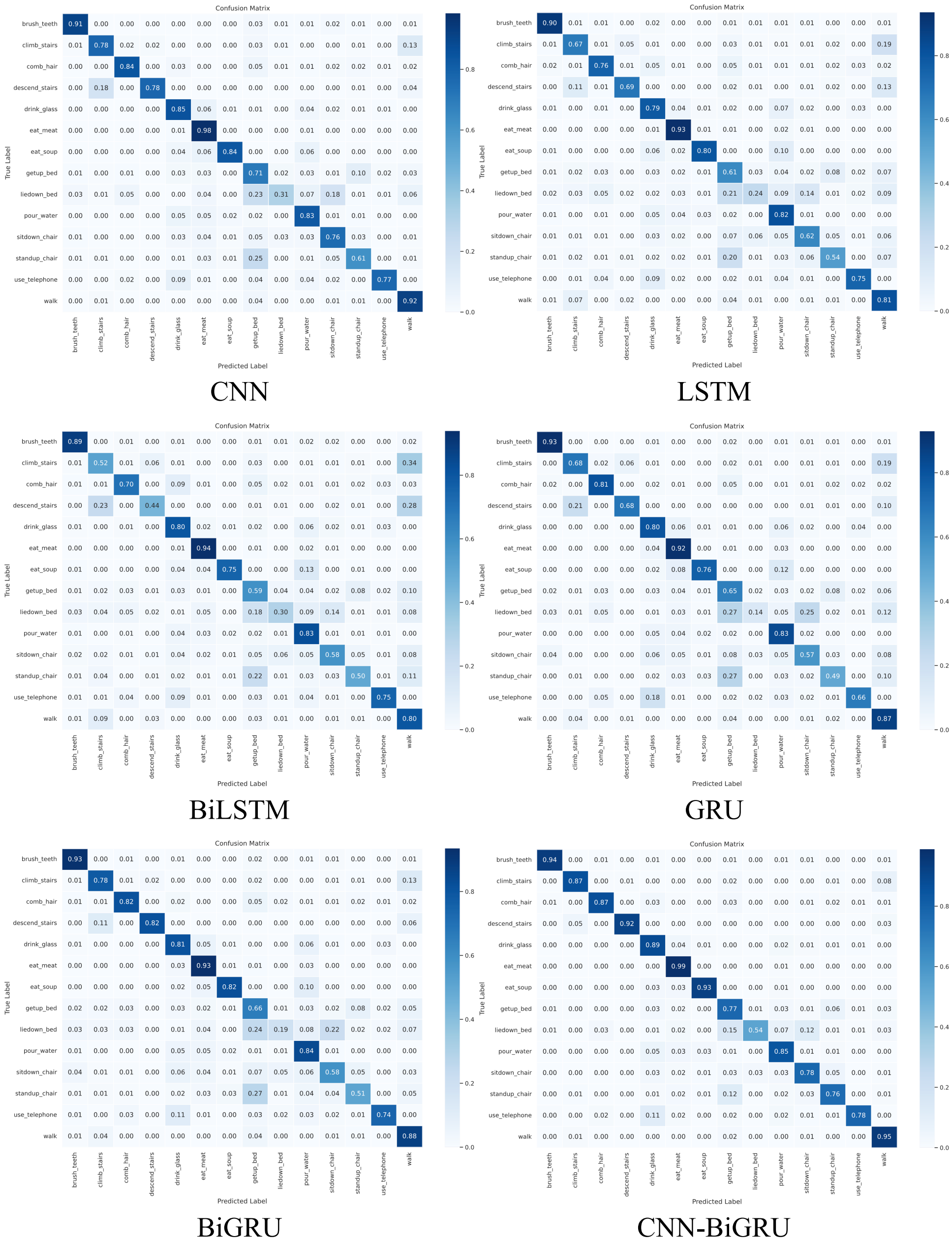

Fig. 3. Confusion matrices of deep learning models used in this study.

extracting spatial features and a BiGRU layer for extracting temporal features. The study employed confusion matrices to analyze the model's performance across different activity types. The results indicate that the proposed model faced challenges in accurately classifying transition activities, such as "Get up from the bed," "Lie down on the bed," and "Sit down on a chair." However, it demonstrated superior performance in manual tasks involving utensils, such as "Brush," "Eat with fork and knife," and "Eat with spoon."

V. Conclusion and Future Works

This research aimed to identify complex activities using wrist-worn sensor data. Five deep learning models, namely CNN, LSTM, BiLSTM, GRU, and BiGRU, were examined, along with the newly proposed CNN-BiGRU, to effectively classify these intricate activities. The evaluation was conducted on the WHARF dataset, a publicly available benchmark for HAR, which includes data from wrist-worn sensors capturing 14 different activities. The results showed that the CNN-BiGRU model achieved the highest accuracy rate of 87.20%.

Future investigations could focus on validating these deep learning models on alternative datasets with a larger number of subjects exhibiting diverse complex patterns. To enhance performance, there is potential for developing low-power and lightweight deep learning networks, as well as implementing innovative data representations based on time-frequency analysis.

Acknowledgment

This research project was supported by Thailand Science Research and Innovation Fund; University of Phayao under

Grant No. FF66-UoE001; National Science, Research and Innovation Fund (NSRF); and King Mongkut's University of Technology North Bangkok with Contract no. KMUTNB-FF-67-B-09.

REFERENCES

[1] B. M. G. Rosa, B. Lo, and E. Yeatman, "Prototype smartwatch device for prolonged physiological monitoring in remote environments," in *2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2022, pp. 1–4.

[2] S. Mekruksavanich, P. Jantawong, N. Hnoohom, and A. Jitpattanakul, "Human activity recognition for people with knee abnormality using surface electromyography and knee angle 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. 483–487.

[3] S. Mekruksavanich and A. Jitpattanakul, "Sport-related activity recognition from wearable sensors using bidirectional gru network," *Intelligent Automation & Soft Computing*, vol. 34, no. 3, pp. 1907–1925, 2022.

[4] L. Zhang, W. Cui, B. Li, Z. Chen, M. Wu, and T. S. Gee, "Privacy-preserving cross-environment human activity recognition," *IEEE Transactions on Cybernetics*, vol. 53, no. 3, pp. 1765–1775, 2023.

[5] W.-Y. Cheng, A. Scotland, F. Lipsmeier, T. Kilchenmann, L. Jin, J. Schjodt-Eriksen, D. Wolf, Y.-P. Zhang-Schaerer, I. F. Garcia, J. Siebourg-Polster, J. Soto, L. Verselis, M. Martin-Facklam, F. Boess, M. Koller, M. Grundman, A. Monsch, R. Postuma, A. Ghosh, T. Kremer, K. Taylor, C. Czech, C. Gossens, and M. Lindemann, "Human activity recognition from sensor-based large-scale continuous monitoring of parkinson's disease patients," in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 2017, pp. 249–250.

[6] S. Mekruksavanich and A. Jitpattanakul, "Deep convolutional neural network with rnns for complex activity recognition using wrist-worn wearable sensor data," *Electronics*, vol. 10, no. 14, 2021.

[7] O. C. Ann and L. B. Theng, "Human activity recognition: A review," in *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, 2014, pp. 389–393.

[8] S. Saini, A. Juneja, and A. Shrivastava, "Human activity recognition using deep learning: Past, present and future," in *2023 1st International Conference on Intelligent Computing and Research Trends (ICRT)*, 2023, pp. 1–6.

[9] S. Mekruksavanich and A. Jitpattanakul, "Deep residual network for smartwatch-based user identification through complex hand movements," *Sensors*, vol. 22, no. 8, 2022.

[10] N. Hnoohom, S. Mekruksavanich, and A. Jitpattanakul, "An efficient resnetse architecture for smoking activity recognition from smartwatch," *Intelligent Automation & Soft Computing*, vol. 35, no. 1, pp. 1245–1259, 2023.

[11] L. Peng, L. Chen, Z. Ye, and Y. Zhang, "Aroma: A deep multi-task learning based simple and complex human activity recognition method using wearable sensors," vol. 2, no. 2, jul 2018.

[12] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors*, vol. 16, no. 4, 2016.

[13] Z. Qin, Y. Zhang, S. Meng, Z. Qin, and K.-K. R. Choo, "Imaging and fusing time series for wearable sensor-based human activity recognition," *Information Fusion*, vol. 53, pp. 80–87, 2020.

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

[15] K. Wang, J. He, and L. Zhang, "Attention-based convolutional neural network for weakly labeled human activities' recognition with wearable sensors," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7598–7604, 2019.

[16] A. Murad and J.-Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors*, vol. 17, no. 11, 2017.

[17] K. Xia, J. Huang, and H. Wang, "Lstm-cnn architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56 855–56 866, 2020.

[18] S. Mekruksavanich, A. Jitpattanakul, K. Sitthithakerngkiet, P. Youplao, and P. Yupapin, "Resnet-se: Channel attention-based deep residual network for complex activity recognition using wrist-worn wearable sensors," *IEEE Access*, vol. 10, pp. 51 142–51 154, 2022.

[19] N. T. Hoai Thu and D. S. Han, "Hihar: A hierarchical hybrid deep learning architecture for wearable sensor-based human activity recognition," *IEEE Access*, vol. 9, pp. 145 271–145 281, 2021.

[20] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, jan 2014.

[21] G. M. Weiss, K. Yoneda, and T. Hayajneh, "Smartphone and smartwatch-based biometrics using activities of daily living," *IEEE Access*, vol. 7, pp. 133 190–133 202, 2019.

[22] S. G. Trost, Y. Zheng, and W.-K. Wong, "Machine learning for activity recognition: hip versus wrist data," *Physiological Measurement*, vol. 35, no. 11, p. 2183, oct 2014.

[23] S. Chernbumroong, A. S. Atkins, and H. Yu, "Activity classification using a single wrist-worn accelerometer," in *2011 5th International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA) Proceedings*, 2011, pp. 1–6.

[24] F. G. da Silva and E. Galeazzo, "Accelerometer based intelligent system for human movement recognition," in *5th IEEE International Workshop on Advances in Sensors and Interfaces IWASI*, 2013, pp. 20–24.

[25] R. I. Ramos-Garcia and A. W. Hoover, "A study of temporal action sequencing during consumption of a meal," in *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics*, ser. BCB'13. New York, NY, USA: Association for Computing Machinery, 2013, p. 68–75.

[26] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, "Detecting periods of eating during free-living by tracking wrist motion," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1253–1260, 2014.

[27] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, "The case for smartwatch-based diet monitoring," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 585–590.

[28] S. Mekruksavanich, P. Jantawong, N. Hnoohom, and A. Jitpattanakul, "Deep learning networks for eating and drinking recognition based on smartwatch sensors," in *2022 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C)*, 2022, pp. 106–111.

[29] S. Mekruksavanich and A. Jitpattanakul, "Cnn-based deep learning network for human activity recognition during physical exercise from accelerometer and photoplethysmographic sensors," *Lecture Notes on Data Engineering and Communications Technologies*, vol. 117, p. 531 – 542, 2022.

[30] K. S. Gajjala and B. Chakraborty, "Human activity recognition based on lstm neural network optimized by pso algorithm," in *2021 IEEE 4th International Conference on Knowledge Innovation and Invention (ICKII)*, 2021, pp. 128–133.

[31] S. Mekruksavanich and A. Jitpattanakul, "Lstm networks using smartphone data for sensor-based human activity recognition in smart homes," *Sensors*, vol. 21, no. 5, p. 1 – 25, 2021.

[32] R. Mutegeki and D. S. Han, "A cnn-lstm approach to human activity recognition," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2020, pp. 362–366.

[33] S. Mekruksavanich and A. Jitpattanakul, "Hybrid convolution neural network with channel attention mechanism for sensor-based human activity recognition," *Scientific Reports*, vol. 13, no. 1, p. 12067, 2023.

[34] N. K. Singh and K. S. Suprabhath, "Har using bi-directional lstm with rnn," in *2021 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, 2021, pp. 153–158.

[35] 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.

[36] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa, "A public domain dataset for adl recognition using wrist-placed accelerometers," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014, pp. 738–743.

[37] D. Dua and C. Graff, "Uci machine learning repository," 2017. [Online]. Available: http://archive.ics.uci.edu/ml