Wearable Sensor Feature Fusion for Human Activity Recognition (HAR): A Proposed Classification Framework

Norfadzlan bin Yusup
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak (UNIMAS)
Kota Samarahan, Sarawak, Malaysia
ynorfadzlan@unimas.my

Adnan Shahid Khan
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak (UNIMAS)
Kota Samarahan, Sarawak, Malaysia
skadnan@unimas.my

Izzatul Nabila bt Sarbini
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak (UNIMAS)
Kota Samarahan, Sarawak, Malaysia
sinabila@unimas.my

Nurul Zawiyah bt Mohamad
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak (UNIMAS)
Kota Samarahan, Sarawak, Malaysia
mnzawiyah@unimas.my

2022 International Conference on Digital Transformation and Intelligence (ICDI) | 979-8-3503-9700-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICDI57181.2022.10007121

Abstract— **Human Activity Recognition (HAR) focuses on detecting people's daily regular activities based on time-series recordings of their actions or motions. Due to the extensive feature engineering and human feature extraction required by traditional machine learning algorithms, they are time-consuming to develop. To identify complicated human behaviors, deep learning approaches are more suited since they can automatically learn the features from the data. In this paper, a feature-fusion concept on handcrafted features and deep learning features is proposed to increase the recognition accuracy of diverse human physical activities using wearable sensors. The deep learning model Long-Short Term Memory based Deep Recurrent Neural Network (LSTM-DRNN) will be used to extract deep features. By fusing the handcrafted produced features with the automatically extracted deep features through the use of deep learning, the performance of the HAR model can be improved, which will result in a greater level of accuracy in the HAR model. Experiments conducted on two publicly available datasets show that the proposed feature fusion achieves a high level of classification accuracy.**

Keywords— Human Activity Recognition (HAR), Deep Learning, Wearable Sensors, Metaheuristic algorithm, Feature Selection

I. INTRODUCTION

Artificial Intelligence (AI) rapid progress has made it feasible to employ HAR in a wide range of study fields because machine learning algorithms enhance the accuracy of human activity recognition, making it viable to apply it in many different research fields. HAR is used mainly as an assistive technology for older adults and healthcare in conjunction with other devices such as the Internet of Things (IoT) with the assistance of sensor or smartphones. Wearable technologies are experiencing growing requirements and interests, especially in healthcare. For example, in defining physical health hazards, the volumes of information produced from wearable devices could be helpful to avoid certain falls or dangers.

In human activity recognition (HAR) models, the accuracy of the performance is heavily reliant on the features that can be extracted from domain expertise. The features are used as input by the classification algorithm to quickly identify physical activities performed by humans. Expertise in the field is required to extract hand-crafted features. As a result, these features are crucial for identifying various human activities. Deep learning methods are used in present-day HAR research applications[1][2]. It has recently become possible to extract features from raw sensory data using deep learning methods. On the other hand, the most recent research in HAR has shown that one cannot disregard the significance of handcrafted features because this information is derived from the domain expertise of specialists [3]. Conventional ways of machine learning extract hand-crafted features that need to be chosen manually. As a result, it is possible that uninformative features may be selected, which will result in an incorrect classification. Deep learning systems are able to automatically extract features that are appropriate for the task at hand and concentrate on improving recognition performance.

II. RELATED WORKS

As deep learning method has progressed, more researchers have started using it for tasks like feature extraction and activity recognition classification. There have been a number of research that have employed sensors to collect data and analyse outcomes in order to construct efficient HAR systems by employing deep-learning algorithms. Based on research by [4], a novel hybrid deep learning model called CNN- GRU was proposed for the classification of complex human activities. In addition, the results were validated using the train, test, and validation datasets. Based on the author’s findings, smartwatches provide a higher level of accuracy than smartphones when it comes to recognising complex human activities. Overall, the findings showed that hybrid deep learning models are capable of efficiently and automatically extracting the spatial-temporal features from raw sensor data in order to classify complex human actions. Furthermore, these models were able to provide better accuracy than other deep learning models that were used in this study and had a relatively complex model architecture. In addition to GRU and CNN, more complicated models of deep neural networks will also be taken into consideration as part of our ongoing research. Because of its potential applications in personalized health and fitness monitoring, the identification of human physical activities has been a topic of active research for quite some time. The accuracy of the performance HAR models is primarily dependent on the features that are derived from domain knowledge. Recently, methods from the field of deep learning have been applied in order to automatically extract features from raw sensory data for use in HAR models.

In the work of [5] the UCI dataset and self-recorded data were used to validate and calibrate one-dimensional convolutional neural network (CNN)-based motion recognition method. Accuracy of 95.99% was achieved when applied to the UCI dataset. The hybrid structure proposed by [6] consists of one CNN layer and one LSTM layer. The dataset provided by UCI was utilised for both training and validation, which resulted in an accuracy of 92.13%. In order to train and validate their model, authors in [7] used the UCI dataset with a bidirectional long short-term memory. Positive results were obtained by using models with three-layer network topologies and no more than 175 units, and the accuracy of these models was 92.67%. To avoid overfitting, authors in [8] developed a five-layer LSTM network topology with L2 regularisation. The UCI data was used for both training and verification, and the resulting accuracy was 93.13%.

A solution has been developed for a 3D skeleton-based HAR problem by combining compact spatio-temporal image encoding with evolutionary algorithm-based feature selection by [9]. The proposed method achieves results that are comparable to those achieved by the existing methods while using fewer parameters. The Ant Lion optimization (ALO) algorithm was used for the feature selection step in this work. In this work, the authors opted to use ALO. In addition to this, we intend to broaden the scope of the experimentation to include numerous types of data (RGB video, IR sequence data, etc.). In the future, we plan to broaden the scope of the experimentation to include additional HAR datasets and possibly even other fields that make use of skeletal data. This will allow us to further improve the robustness of the framework.

In the study by [10] the authors proposed a new deep learning model for device-free human activity recognition, with the crucial innovative CSI-Correlation Feature Extraction (CCFE) method for data preprocessing as the centerpiece of the approach. The CCFE method makes use of a recursive algorithm to improve activity-related signals and then computes the correlation across segmented signals in both the time and frequency domains in order to condense the improved signals. This allows the improved signals to be transmitted in a more compact form. When compared to traditional Recurrent Neural Network (RNN)-based sensing methods, the proposed method provides noticeably enhanced recognition accuracy, while at the same time greatly reducing the amount of training time required. The comparison is based on comprehensive experimental results.

Deep learning methods are used in present-day HAR research applications. Deep learning can automatically learn from raw sensor data that are less available and refine layer-by-layer parameters based on the assumption that the decoded output matches inputs. Deep learning will become increasingly relevant in the future in various application fields, e.g. recognising adult behaviour, child activity and smart home automation. Despite its effectiveness in differentiating between inter-class similarities activities behaviour, further improvements are still required before implementing the model. The Deep Convolutional Neural Network was used by [11] to recognise the operation from accelerometer sensor. The reported result showed the related output and obtained an almost perfect classification for differentiating different walking movements. However, SVM classifier has outperformed the proposed deep learning algorithm for classification in terms classification of static activities. Deep learning is often associated with higher levels of complexity and ambiguity, making it unfeasible for real-time applications[12][13][14]. CNN, however, has successfully distinguished between distinct stair-based activities, but overall stair accuracy is still unsatisfactory[11].

III. HAR CLASSIFICATION PROCESS

There are a few stages that need to be completed in order to classify HAR data, and they are as follows:

1. Preprocessing
   Pre-processing is essential in order to ensure that the signal is interpreted sufficiently so that the classification model can provide a high degree of prediction accuracy. In this stage, unnecessary or unimportant information can be removed before any further procedure is performed. Unnecessary details, missing features or instances, and poorly defined behaviours are assumed to be prone to noise. Furthermore, during data collection, a device may have been accidentally dropped or misplaced. It will then generate an unnecessary signal that affects signal representation. Figure 1 below shows the example of raw data of static and dynamic activities (Sit, Stand, Walk, Walking Up, Walking Down) captured continuously using iPhone 13 Max based on accelerometer and angular velocity.

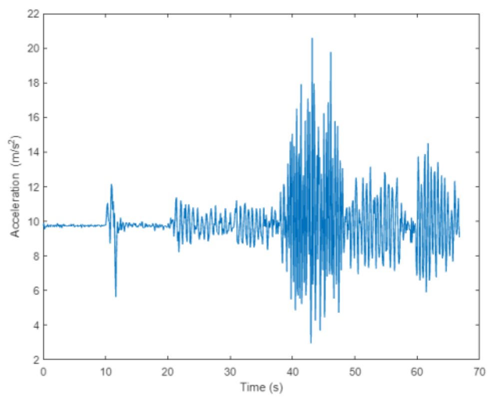

Fig. 1. Raw HAR data

2. Segmentation

Segmentation is necessary in order to interpret the signal signature in a more meaningful manner. Therefore, because of its less complexity and the fact that the continuous signal is useful, a sliding window segmentation is implemented. Every acceleration (*x, y, z*) and angular velocity (*x, y, z*) data stream is separated into the same size as the predefined window segment. A sliding window with a 50% gap between two consecutive window segments is normally used to minimize potential errors induced by transition state noise. Furthermore, additional features for defining the activities attribute are derived from each section of the windows. Extracted features will reflect the relevant information by reducing the data stream representation. The extracted feature subsets will then be retained as a classification model input variable.

3. Feature Extraction

Features are required as an input for the classifiers in order for classification to work. Handcrafted features such as the median, min, and median frequency measurement measures in the time domain or the frequency domain from each window field to record proper data representation to distinguish between HAR activities. Features are taken from the simple calculation of statistical analysis in each window segment; the correlation between different window segments is calculated. Time-domain features are simple and directly derived such that they are mostly added to behaviour recognition. Frequency domain features are often taken into consideration to enhance accuracy efficiency. These features define the sum of signal energy that is transmitted across the frequency spectrum by capturing the repeated existence of the signal. Table 1 below shows the handcrafted time-domain and frequency-domain features commonly extracted in HAR datasets.

TABLE 1. TIME AND FREQUENCY DOMAIN FEATURES

Feature Domain	Feature
Time-Domain Features	Mean value
	Standard deviation
	Median absolute deviation
	Largest value in array
	Smallest value in array
	Auto regression coefficients with Burg order equal to 4
	Signal Magnitude Area
Frequency-Domain Features	Weighted average of the frequency of signals
	Skewness of the frequency domain signal
	Kurtosis of the frequency domain signal
	Spectral energy of frequency band
	Index of the frequency component with largest magnitude

4. Feature Selection and Feature Reduction

The number of features observed has risen from dozens to thousands in several machine learning implementations. However, the uncertainty of the learning method also gets compounded by the inclusion of trivial and irrelevant features. The purpose is, therefore, to choose the best feature subsets in the classifier model before applying the reduced feature sub-sets. The selection of features relies on various factors. Features are able to effectively identify the performance class and, at the same time, minimise redundant features. Consequently, the selection of features is vital to boost the model accuracy.

5. Classification

The standard method to determine the importance of classification is to split the feature subsets into two separate subsets: training and testing. Patterns are obtained from a number of samples or instances in the training group. Test subset is used to verify the classifier model output by testing the classifier model's ability to generalize to identify the instance that has never been seen. A validation technique approach is required in such circumstances to calculate the level of results for both training and test subsets. Cross-validation is the most used approach known for calculating prediction error.

Generally, cross-validation entails methods for splitting the sample into multiple training and testing data sets. The *k*-fold cross-validation, hold-out, and leave-one-out validation method are the most common techniques for different classification problems. Performance measures or metrics are essential for the evaluation of recognition performance in any classification problem. Average accuracy is commonly used for the calculation of average recognition performance, but unbalanced class distribution is not efficiently measurable. For example, if the class ratio is 10:90, 90 percent accuracy could be obtained even if all instances are divided into a majority class. Therefore, several other metrics such as precision, recall and F-measurement are used. In addition, statistical tool such as Matthew Correlation Coefficient is usually used to measure the quality of the classification model.

IV. Proposed Solutions

The deep learning model Long-Short Term Memory based Deep Recurrent Neural Network (LSTM-DRNN) will be used to extract features and the deep features extracted will be fed to the Binary Grey Wolf Algorithm (BWGO) to reduce features and classification. Recent research has revealed that the performance of neural networks in deep learning approaches is superior to that of conventional machine learning methods [15], [16]. Among the deep learning approaches, a Recurrent Neural Network (RNN) architecture that makes use of Long Short-Term Memory (LSTM) demonstrates superior performance in a variety of research domains. The RNN variety of neural networks is capable of processing sequential data, and the LSTM is one subset of the RNN family. Figure 2 below show the Unidirectional LSTM-DRNN model.

The Grey Wolf Optimization (GWO) algorithm, introduced by [17] is one of the most recent bio-inspired optimization approaches that simulates the hunting activity that grey wolves engage in while they are in their natural habitat. There are two methods in Binary Grey Wolf Optimization (BGWO) algorithm as proposed by [18]. In the first method, the individual steps that lead up to the first three best solutions are binarized.

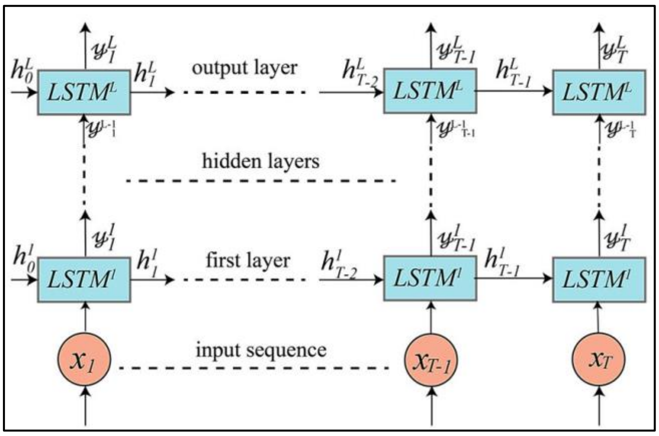

Fig. 2. LSTM-DRNN Model [16]

After that, stochastic crossover is carried out between the three fundamental moves in order to locate the most recent binary grey wolf location. In the second method, a sigmoidal function is first employed to compress the continuously updated location, and then stochastically thresholded values are utilised to determine the most recent binary position of the grey wolf. In the feature selection domain, the two approaches for binary grey wolf are used to discover a feature subset that maximises classification accuracy while lowering the amount of features that are selected.

The evaluation of the position of the prey as seen by α wolves, β wolves, and δ wolves is an important step in the optimization process for the GWO algorithm. The remaining wolves will use this location as a guideline for determining where they are in relation to the prey, and they will do so in a random fashion[19]. Figure 3 depicts the process by which grey wolves update their positions depending on the location information of α wolf, β wolf, and δ wolf. This process takes place based on the information provided by wolves α, β, and δ. The evaluation of the position of the prey as seen by α wolves, β wolves, and δ wolves is an important step in the optimization process for the GWO algorithm. The remaining wolves will use this location as a guideline for determining where they are in relation to the prey, and they will do so in a random fashion. Figure 3 depicts the process by which grey wolves update their positions depending on the location information of α wolf, β wolf, and δ wolf. This process takes place based on the information provided by wolves α, β, and δ.

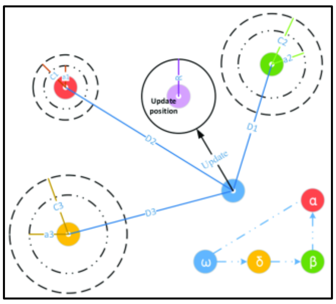

Fig. 3. Grey Wolf Optimization Process [20]

In general, the proposed classification for activity recognition is divided into four main stages: pre-processing stage, feature extraction stage, feature selection stage, and classification stage as illustrated in Figure 4. Initially, the recorded acceleration and angular velocity data stream undergoes the filtering process to remove the unwanted information (high frequency component) from body acceleration. Afterwards, filtered body acceleration is divided into several sizes of window segments. Next, features are extracted from each window segment before it is fed for the next process. In such situations, some features were derived from time-domain and frequency-domain features (hand-crafted features).

Many of the drawbacks associated with previous iterative machine learning approaches have been addressed using Deep Learning methods. Deep Learning approaches that can automatically extract features are better suited for increasingly challenging problems.

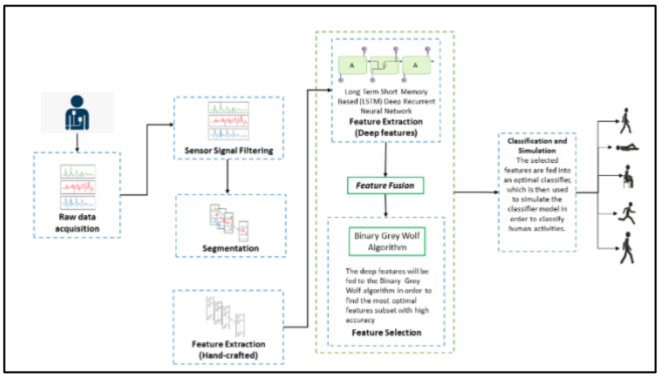

Fig. 4. Proposed HAR Classification framework

When dealing with time-based data, Recurrent Neural Networks are a good fit since they have a sequential component. Features may be retrieved using LSTM-DRNN. The HAR model is significantly impacted by both hand-crafted and deep learned features. For this reason, a feature-fusion idea is suggested to increase the recognition accuracy of diverse human physical activities using wearable sensors. Subsequently, the Binary Grey Wolf algorithm is developed to reduce the feature dimension by selecting the most relevant features to be classified. Lastly, the quality of the selected feature subsets is figured out by using the classifier model. Ensemble Random Forest will be employed as base classifier to deal with the problem of distinguishing activities.

This research addressed the problem of extracting and selecting features for HAR classification. We proposed a feature-fusion concept to increase the recognition accuracy of diverse human physical activities using wearable sensors. The deep learning model Long-Short Term Memory based Deep Recurrent Neural Network (LSTM-DRNN) will be used to extract features and the deep features extracted will be fed to the Binary Grey Wolf Algorithm (BWGO) to select relevant features and classification. Subsequently, the Binary Grey Wolf algorithm will be developed to reduce the feature dimension by selecting the most relevant features to be classified.

V. EXPERIMENTAL AND RESULTS

In this research, two publicly available physical activities accelerometer sensor datasets are utilized: SBHAR and USC-HAD. Table 2 briefly described the general information of the datasets. Both datasets in this research employed accelerometers and gyroscope sensors, which denotes as (A_x, A_y, A_z) and. (G_x, G_y, G_z) where A and G are referring to Accelerometer and Gyroscope, while *x,y,z* referring to the three-dimensional (3D) axes of both sensors.

TABLE 2. DATASET

Dataset	Wearable sensors	Sensor Position	No of Activity	No. of Subject
SBHAR [21]	Accelerometers (A_x, A_y, A_z)	Waist	6	30
USC-HAD [22]	Gyroscope (G_x, G_y, G_z)		10	14

For SBHAR Dataset, a constant rate of 50 Hz is captured for each operation at three-axial linear accelerations (accelerometer) and three-axial angular velocity (gyroscopes). Data labelling is performed manually with a video taken each time during the activity. SBHAR is mobile data collection based on human behaviour identification. The subjects' movements include walking, upstairs walking, walking downstairs and upstairs, laying, sitting, and standing. For USC-HAD dataset, the sensing unit MotionNode was mounted on the subject's upper thigh for 14 subjects, and the frequency of sampling used is 100Hz. This device is embedded with three-axial accelerometer and three-axial gyroscope sensors. The subjects were carrying out 12 different styles of activities: walking forward, walking left, walking right and upstairs, walking downwards, walking up and down, sitting, standing, sleeping, elevator up and downwards. There is a total of six activities for the SBHAR dataset and ten activities for the USC-HAD dataset employed in this research, with the number of subjects 30 and 14, respectively. In the experiment, 70% of the data is designated as training data, 30% is designated as testing data. The model is evaluated based on accuracy and F1-Score in (1) and (2) where the number of correctly recognized class (TP), the number of correctly recognized examples that do not belong to the class (TN), and the number of examples that were neither correctly assigned class (FP) nor recognized as the class (FN) will be computed.

$$Accuracy = \frac{tp+tn}{tp+fn+fp+tn} \quad (1)$$

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (2)$$

As can be seen in Figure 2, the model is constructed using DRNN based on a unidirectional LSTM. In both datasets, the best results are achieved by using a DRNN with four layers. The first layer receives an input sequence of samples (x1, x2,..., xT) at time t (t = 1, 2,..., T). To train the deep learning model, Adam optimizer is employed and TensorFlow was used for training. Adam is an optimization algorithm that updates model parameters and backpropagates the gradient to minimise the cost function. In the experiment, the dropout method was employed to prevent overfitting. It is common practise to apply dropout to all network nodes, but this practise is reserved for interlayer links (not on recurrent-connections or intra-cell connections).

The HAR model is significantly affected by both manually extracted features and features learned by Deep Learning algorithms. Our goal is to increase the identification accuracy of different human physical activities using accelerometers and gyroscopes, so a feature fusion concept was proposed to make use of both features. LSTM model of DRNN was investigated for use in feature learning throughout the course of this research. To train features, the raw sensory data is fed in. The fully connected layer waits until the end of the process before taking the learned features and using them as input. At the same time, another fully connected layer uses the handcrafted features in Table 1 an as input to derive additional conceptual features. At last, the two features are combined and fed into the softmax layer for human physical activity classification.

TABLE 3. FEATURE GROUP PERFORMANCE

Dataset	Features	F1-Score	Classification Accuracy
SBHAR	Time-Domain	95.2%	95.4%
	Frequency-Domain	93.2%	93.4%
	DRNN	96.6%	96.7%
	Proposed Feature Fusion	98.1%	98.2%
USC-HAD	Time-Domain	96.4%	96.6%
	Frequency-Domain	93.3%	93.5%
	DRNN	97.6%	97.8%
	Proposed Feature Fusion	98.5%	98.6%

The accuracy of feature fusion classification shows a better accuracy of 98.2% compared to the time-domain features group with 97.4% and the frequency-domain features group with 93.4% in SBHAR dataset. In USC-HAD dataset, the accuracy of the feature fusion classification shows a better accuracy of 98.6% compared to the time-domain features group with 96.6% and the frequency-domain features group with 93.5%. The results of the experiment show that the proposed feature fusion approach can properly leverage the complementarity between the group features and improve the efficiency of the classification.

TABLE 4. CLASSIFIFCATION ACCURACY RESULTS

Dataset	Classification Accuracy			
	SVM	CNN	DRNN	Proposed Feature Fusion
SBHAR	96.10%	95.2%	96.7%	98.2%
USC-HAD	95.99%	97%	97.8%	98.6%

According to Table 4, the proposed feature fusion with LSTM-DRNN for the SBHAR dataset exhibits a high accuracy of 98.2% when compared to the other techniques such as DRNN, CNN, and SVM. On the USC-HAD dataset, the proposed feature fusion obtained 98.6% accuracy, outperforming existing techniques such as DRNN, CNN, and SVM.

VI. Conclusion

In this research, we presented a proposed feature fusion model that utilises both machine-generated features and LSTM-DRNN. A fully layered fusion model was developed by concatenating the fully layered models. There were two publicly available datasets used in HAR's implementation of this model. The accuracy of classification in the SBHAR dataset was 98.2%, whereas the accuracy of classification in the USC-HAD data was 98.6%. A variety of complex motions, not just sitting, standing, and walking, will be recognised as physical activities in the near future. Because of this, it is reasonable to anticipate that additional motions or actions will be involved in complex processes. Complicated activities like drinking coffee, cleaning, and playing certain sports like badminton, for example, require the subject to make bodily movements in order to execute each activity using various body parts. Therefore, features of the activity that set it apart from others are necessary for accurate classification of the activity. The features need to be unique while also having a strong correlation to the type of class activities. The selection of the incorrect features can lead to the production of inaccurate results. As part of our ongoing research, we are going to apply the proposed feature fusion to the data that we have collected on our own. Aside from that, the BGWO algorithm will be employed so that additional features can be optimised and a high level of classification accuracy can be achieved.

Acknowledgment

This research work is fully funded by Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak.

References

[1] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, and J. S. Suri, "Human activity recognition in artificial intelligence framework: a narrative review," *Artif. Intell. Rev.*, vol. 55, no. 6, pp. 4755–4808, 2022, doi: 10.1007/s10462-021-10116-x.

[2] R. Abdel-Salam, R. Mostafa, and M. Hadhood, "Human Activity Recognition Using Wearable Sensors: Review, Challenges, Evaluation Benchmark BT - Deep Learning for Human Activity Recognition," 2021, pp. 1–15.

[3] C. T. Yen, J. X. Liao, and Y. K. Huang, "Feature fusion of a deep-learning algorithm into wearable sensor devices for human activity recognition," *Sensors*, vol. 21, no. 24, 2021, doi: 10.3390/s21248294.

[4] S. Gupta, "Deep learning based human activity recognition (HAR) using wearable sensor data," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100046, 2021, doi: 10.1016/j.jjimei.2021.100046.

[5] C.-T. Yen, J.-X. Liao, and Y.-K. Huang, "Human daily activity recognition performed using wearable inertial sensors combined with deep learning algorithms," *Ieee Access*, vol. 8, pp. 174105–174114, 2020.

[6] R. Mutegeki and D. S. Han, "A CNN-LSTM Approach to Human Activity Recognition," *2020 Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2020*, pp. 362–366, 2020, doi: 10.1109/ICAIIC48513.2020.9065078.

[7] F. Hernández, L. F. Suárez, J. Villamizar, and M. Altuve, "Human activity recognition on smartphones using a bidirectional LSTM network," in *2019 XXII symposium on image, signal processing and artificial vision (STSIVA)*, 2019, pp. 1–5.

[8] M. Ullah, H. Ullah, S. D. Khan, and F. A. Cheikh, "Stacked lstm network for human activity recognition using smartphone data," in *2019 8th European workshop on visual information processing (EUVIP)*, 2019, pp. 175–180.

[9] H. Basak, R. Kundu, P. K. Singh, M. F. Ijaz, M. Woźniak, and R. Sarkar, "A union of deep learning and swarm-based optimization for 3D human action recognition," *Sci. Rep.*, vol. 12, no. 1, pp. 1–17, 2022, doi: 10.1038/s41598-022-09293-8.

[10] Z. Shi, J. A. Zhang, R. Xu, and Q. Cheng, "Deep Learning Networks for Human Activity Recognition with CSI Correlation Feature Extraction," *IEEE Int. Conf. Commun.*, vol. 2019-May, 2019, doi: 10.1109/ICC.2019.8761445.

[11] C. A. Ronao and S. B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.*, vol. 59, pp. 235–244, 2016, doi: 10.1016/j.eswa.2016.04.032.

[12] S. Kalogiannis, K. Deltouzos, and J. Ellul, "Recognizing Physical Activity of Older People from Wearable Sensors and Inconsistent Data," pp. 1–19, 2019, doi: 10.3390/s19040880.

[13] W. Sousa Lima, H. L. de Souza Bragança, K. G. Montero Quispe, and E. J. Pereira Souto, "Human activity recognition based on symbolic representation algorithms for inertial sensors," *Sensors (Switzerland)*, vol. 18, no. 11, 2018, doi: 10.3390/s18114045.

[14] N. Hnoohom, A. Jitpattanakul, and S. Mekruksavanich, "Real-life Human Activity Recognition with Tri-axial Accelerometer Data from Smartphone using Hybrid Long Short-Term Memory Networks," *Proc. - 2020 15th Int. Jt. Symp. Artif. Intell. Nat. Lang. Process. iSAI-NLP 2020*, 2020, doi: 10.1109/iSAI-NLP51646.2020.9376839.

[15] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of LSTM," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 6, pp. 1307–1317, 2020, doi: 10.1007/s13042-019-01041-1.

[16] A. Murad and J. Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors (Switzerland)*, vol. 17, no. 11, 2017, doi: 10.3390/s17112556.

[17] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: https://doi.org/10.1016/j.advengsoft.2013.12.007.

[18] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371–381, 2016, doi: 10.1016/j.neucom.2015.06.083.

[19] G. Zhao, T. Cao, Y. Wang, H. Zhou, C. Zhang, and C. Wan, "Optimal Sizing of Isolated Microgrid Containing Photovoltaic/Photothermal/Wind/Diesel/Battery," *Int. J. Photoenergy*, vol. 2021, p. 5566597, 2021, doi: 10.1155/2021/5566597.

[20] S. Dai, D. Niu, and Y. Han, "Forecasting of power grid investment in China based on support vector machine optimized by differential evolution algorithm and greywolf optimization algorithm," *Appl. Sci.*, vol. 8, no. 4, 2018, doi: 10.3390/app8040636.

[21] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," *ESANN 2013 proceedings, 21st Eur. Symp. Artif. Neural Networks, Comput. Intell. Mach. Learn.*, no. April, pp. 437–442, 2013.

[22] M. Zhang and A. A. Sawchuk, "USC-HAD: A Daily Activity Dataset for Ubiquitous Activity Recognition Using Wearable Sensors," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 1036–1043, doi: 10.1145/2370216.2370438.