Transactions on Data Analysis in Social Science 6(4) (2024) 222-230

ISSN Online: 2821-1936

Transactions on Data Analysis in Social Science

Journal Homepage: https://transoscience.ir

Human Activity Recognition Based on Deep Learning Using Sensor Data

S. Mohammadi[1], S. M. Anisheh[2,*]

[1] Department of Education, District 2, Sari, Iran
[2] Department of Electrical Engineering, Hadaf Higher Education Institute, Sari, Iran

ARTICLE INFO

Article History:
Received 3 July 2024
Received in revised form 20 September 2024
Accepted 28 November 2024
Available online 5 December 2024

Keywords:
Human Activity Recognition, Deep Learning, Convolutional Neural Networks, Particle Swarm Optimization Algorithm.

ABSTRACT

Human Activity Recognition (HAR) refers to the process of detecting, identifying, and classifying human activities from sensor data through the application of artificial intelligence techniques. Over the past few decades, HAR has emerged as a rapidly evolving research domain with significant implications in various fields, including video surveillance, identity authentication, smart home automation, healthcare monitoring, and human–computer interaction. In particular, within surveillance systems, timely and accurate recognition of human activities can serve as a preventive measure against incidents such as theft, vandalism, or other suspicious behaviors, thereby enhancing public safety. Among various AI-based approaches, deep learning models especially Convolutional Neural Networks (CNNs) have shown remarkable capabilities in automatically extracting high-level spatiotemporal features and achieving robust classification performance. However, CNN architectures often require extensive hyperparameter tuning to maximize their accuracy and efficiency. To address this challenge, the present study proposes an enhanced CNN model whose parameters are optimized using the Particle Swarm Optimization (PSO) algorithm. The PSO-driven optimization process aims to improve feature extraction quality, reduce overfitting, and enhance generalization capabilities. The proposed framework is implemented and experimentally evaluated on the Wiezmann dataset. Comparative analysis demonstrates that our approach achieves superior recognition accuracy and computational efficiency compared to several existing state-of-the-art methods.

1. INTRODUCTION

Human Activity Recognition (HAR) refers to the process of identifying and classifying human activities using data collected from sensors such as accelerometers, gyroscopes, cameras, and other wearable devices. This process is especially applicable in fields such as health monitoring, smart home automation, and security. HAR plays a significant role in enhancing daily life, monitoring patients, detecting unauthorized activities, and improving safety

* Corresponding Author: s.m.anisheh@gmail.com
Department of Electrical Engineering, Hadaf Higher Education Institute, Sari, Iran

http://dx.doi.org/10.47176/TDASS.2024.222

across various environments. This technology enables humans to interact more intelligently and efficiently with devices and systems [1,2].

Key challenges in HAR include [2]:

1. **Sensor Noise**: Sensors may collect incomplete or inaccurate data.
2. **Activity Diversity**: Human activities are highly varied, making accurate classification difficult.
3. **Scalability**: Adapting systems to large and continuously growing datasets is time-consuming.
4. **Power Consumption**: Real-time processing demands significant computational resources.
5. **Data Ambiguity**: Similar activities may produce overlapping patterns.

Despite these challenges, deep learning offers notable advantages [2]:

1. **Automatic Feature Extraction**: Deep learning models can extract features automatically without manual engineering.
2. **High Accuracy**: Deep models, such as Convolutional Neural Networks (CNNs), perform exceptionally well in recognizing complex activities.
3. **Flexibility**: These models can be easily trained for various activity types.
4. **Multi-Modal Processing**: They can integrate data from multiple sensor types.
5. **Adaptability**: They improve performance with exposure to more and diverse data.

2. RELATED WORK

Human Activity Recognition (HAR) is a critical area of research in ubiquitous computing, with applications in healthcare, smart homes, sports, and ambient-assisted living. HAR aims to identify and classify human activities, such as walking, running, or sitting, using data collected from various sensors, including wearable devices, smartphones, and inertial measurement units (IMUs). The advent of deep learning has significantly advanced HAR by enabling the automatic extraction of complex features from raw sensor data, overcoming limitations of traditional machine learning approaches that rely heavily on handcrafted features [3].

Early HAR systems primarily utilized wearable sensors, such as accelerometers and gyroscopes, to capture motion data. These systems employed statistical and machine learning techniques, such as Support Vector Machines (SVMs) and decision trees, to classify activities [4]. However, these methods often struggled with the high variability and noise inherent in sensor data, necessitating manual feature engineering, which is time-consuming and domain-specific [5]. The introduction of deep learning techniques, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), has revolutionized HAR by enabling end-to-end learning, where raw sensor data is directly processed to produce activity classifications [6].

CNNs have proven effective in HAR due to their ability to capture spatial patterns in time-series sensor data. For instance, studies have demonstrated that CNNs can extract robust features from accelerometer and gyroscope signals, improving recognition accuracy for activities like walking and climbing stairs [7]. Similarly, Long Short-Term Memory (LSTM) networks, a type of RNN, are well-suited for modeling temporal dependencies in sequential data, making them ideal for continuous activity recognition tasks [8]. Combining CNNs and LSTMs has further enhanced performance by leveraging both spatial and temporal features, particularly for multimodal sensor data from wearables and smartphones [9].

Smartphones and wrist-worn devices have become popular platforms for HAR due to their widespread availability and embedded sensors. Research has shown that deep learning models trained on smartphone sensor data can achieve high accuracy in recognizing complex activities, such as eating or typing, even in naturalistic settings [10]. However, challenges remain, including handling noisy sensor data caused by device placement, user variability, or environmental factors. Robust HAR systems have been developed to address these issues by incorporating noise-tolerant deep learning architectures, such as autoencoders and attention mechanisms [11].

In smart home applications, HAR systems integrate data from multiple sensors, such as IMUs and ambient sensors, to monitor activities of daily living (ADLs) for elderly care or health monitoring. Graph Convolutional Networks (GCNs) have emerged as a promising approach for modeling relationships between multiple sensor streams in such environments [12]. Additionally, transfer learning has been explored to improve HAR model generalizability across different users and devices, reducing the need for extensive labeled datasets [13].

Despite these advances, HAR faces several challenges. Overlapping activities, where multiple actions occur simultaneously, pose difficulties for accurate classification. Recent studies have proposed continuous detection frameworks to address this issue, achieving promising results in distinguishing actions of interest among overlapping activities [14]. Another challenge is the computational complexity of deep learning models, which can be prohibitive for resource-constrained devices like wearables. Lightweight models and optimization techniques, such as quantization and pruning, have been developed to address this limitation [15].

Ensemble learning has also gained attention in HAR, combining multiple deep learning models to improve robustness and accuracy. For example, ensemble methods integrating CNNs and LSTMs have shown superior performance in recognizing complex activities compared to single-model approaches [16]. Furthermore, particle swarm optimization (PSO) has been applied to optimize hyperparameters of deep learning models for HAR, enhancing model performance and convergence speed [17].

In conclusion, deep learning has significantly advanced HAR by leveraging sensor data from wearables, smartphones, and ambient systems. While challenges such as noisy data, overlapping activities, and computational constraints persist, ongoing research continues to address these issues through innovative architectures and optimization techniques. Future work should focus on developing scalable, robust, and energy-efficient HAR systems to enable real-world deployment in diverse applications.

3. PROPOSED METHOD

The use of Convolutional Neural Networks (CNNs) in HAR has become increasingly important due to their superior ability to process complex data, reduce errors, and enhance both speed and accuracy. With growing access to large-scale data and advances in hardware, CNNs continue to expand in applications across industries, healthcare, and security.

3.1. Deep Learning

Deep learning, a subset of machine learning, involves processing data using multilayered artificial neural networks. Inspired by biological neural networks, it is designed to solve complex problems using large datasets. Deep learning uses these multilayered networks to automatically extract and learn meaningful features from raw data. With increased computational power and the availability of massive datasets, deep learning has become a key tool in fields such as artificial intelligence, image processing, computer vision, natural language processing, and beyond [18].

CNNs, a specialized form of deep learning, are particularly effective for processing structured multidimensional data such as images and videos. These models utilize convolutional operations to extract features and reduce data dimensions, yielding remarkable accuracy in object detection, facial recognition, and medical image analysis.

The advantages of using CNNs in HAR include:

High-Accuracy Feature Extraction: CNNs automatically and optimally extract features from raw sensor and image data. Unlike traditional methods that require manual feature design, CNNs can layer-by-layer detect complex and meaningful patterns, thereby improving recognition accuracy.

Handling Complex, Multidimensional Data: HAR data, such as images or sensor streams, are often complex and high-dimensional. CNNs are inherently designed for such data and are well-suited for multimodal inputs.

Versatile Applications: CNN-based HAR has been successfully applied in health monitoring, security, sports, and smart home automation. For instance, in healthcare, CNNs can support patient monitoring and early disease detection.

Reduced Dependency on Manual Feature Engineering: CNNs significantly lessen the need for hand-crafted features, which is vital when data is complex or domain knowledge is limited.

High Processing Speed: Thanks to their parallel structure and compatibility with hardware accelerators like GPUs, CNNs can process large datasets rapidly an essential feature for real-time applications.

Noise and Environmental Robustness: CNNs are naturally more resilient to noise and environmental changes, enabling accurate activity recognition under varied conditions.

Transfer Learning Capability: CNNs are well-suited to transfer learning, allowing models trained on large datasets to be fine-tuned for smaller datasets helpful when labeled data is scarce.

Recognition of Complex Activities: CNNs can differentiate between similar or contextually overlapping activities, boosting performance in surveillance and behavior monitoring systems.

Key layers in a CNN include:

- **Convolution Layer**: Extracts local patterns such as edges or textures.
- **Activation Layer (e.g., ReLU)**: Adds non-linearity to the network.
- **Pooling Layer**: Reduces spatial dimensions and retains key features.
- **Fully Connected Layer**: Responsible for the final classification.

3.2. Particle Swarm Optimization (PSO)

The Particle Swarm Optimization (PSO) algorithm is a metaheuristic technique designed to find near-optimal solutions in complex optimization problems. It belongs to the family of global search algorithms and is inspired by the collective behavior of natural swarms, such as birds or fish moving in search of food.

PSO utilizes a population of particles, each representing a candidate solution. These particles move collectively through the search space, updating their positions to approach the optimal solution.

PSO algorithm steps:

Initialization: Randomly generate an initial population of particles.

Fitness Evaluation: Compute the fitness (objective value) of each particle.

Update pBest and gBest: Update the personal best (pBest) and global best (gBest) positions.

Velocity and Position Update: Adjust the velocity and position of each particle.

Iteration: Repeat the process until a stopping criterion is met (e.g., max iterations or desired accuracy).

In this study, the PSO algorithm is used to optimize the parameters of the CNN model [16–17]. Design variables include filter size, number of filters, and stride size. These parameters are fine-tuned using PSO to maximize the objective function, which is the activity recognition accuracy.

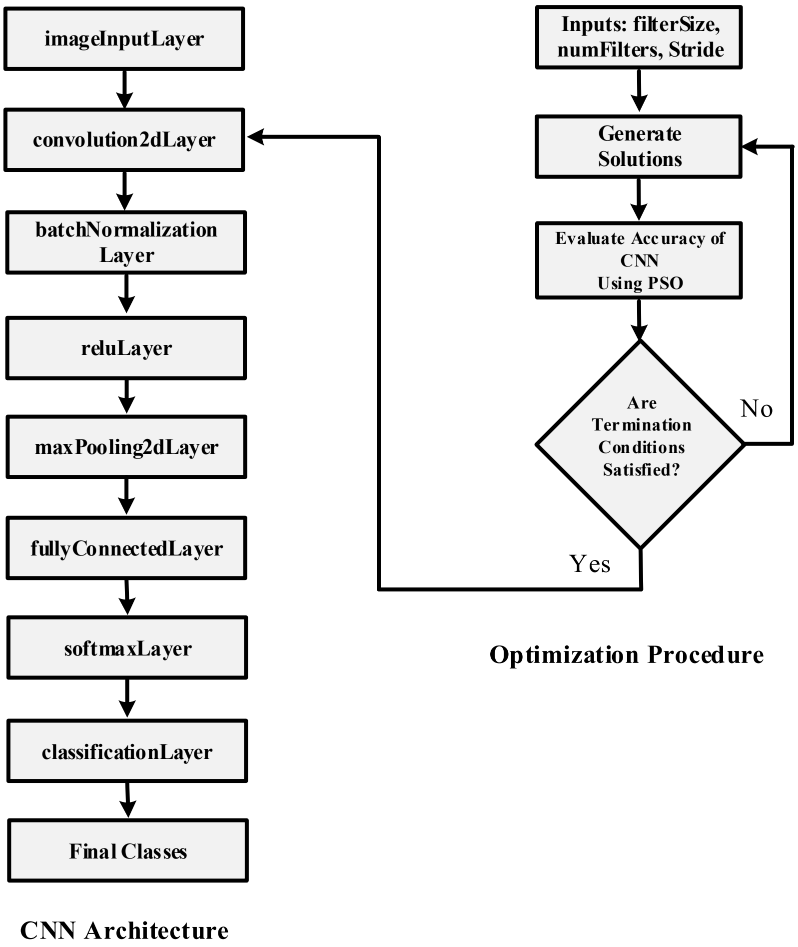

Fig. 1. Flowchart of the Proposed Method.

4. SIMULATION RESULTS

The proposed method and the existing approach were implemented using MATLAB software developed by MathWorks. The simulations were executed on a system equipped with an Intel(R) Core(TM) i5-4460 CPU @ 3.2 GHz and 16 GB of RAM. The parameters of the Particle Swarm Optimization (PSO) algorithm are presented in Table 1.

Table 1. Parameters of the Particle Swarm Optimization Algorithm

Variable	Value
Cognitive coefficient (c1)	2
Social coefficient (c2)	2
Number of particles	100
Number of iterations	1000
Minimum inertia weight	0.4
Maximum inertia weight	0.9
Dimensions	3

These parameters were selected to strike a balance between the speed and accuracy of the algorithm. The Weizmann dataset is one of the most well-known datasets in the field of machine learning and image processing, particularly used for research on human activity recognition in images and videos. This dataset is specifically designed for the identification and modeling of human activities in visual data. The Weizmann dataset comprises images and video sequences of human activities, collected by the Department of Computer Science at the Weizmann Institute of Science. It includes ten different types of human activities, each performed by nine different individuals. An example of the images from the Weizmann dataset is shown in Figure 2 [18].

Fig. 2. Sample images from the Weizmann dataset.

Using the trained network, the labels of the test data are determined and the final test accuracy is estimated. In the case of the Weizmann dataset, the predicted labels match the actual test set labels with an accuracy of 98%. A comparison of the accuracy of the proposed method with existing methods on different datasets is presented in Figure 3. The results demonstrate the high efficiency of the proposed method (highlighted in red). The numbers on the horizontal axis correspond to the proposed method and the existing references.

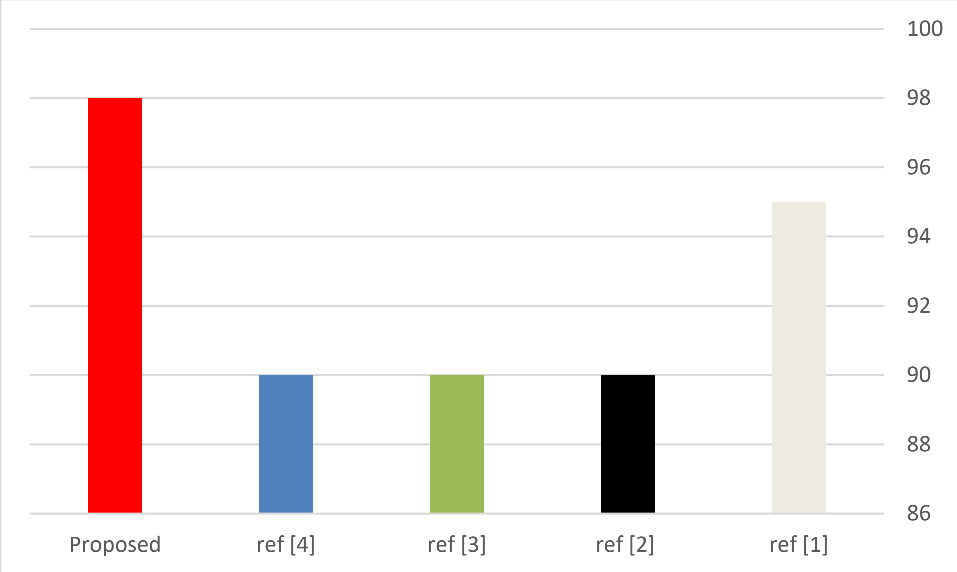

Fig. 3. Comparison of the accuracy of the proposed method with existing methods.

5. CONCLUSION

Convolutional Neural Networks (CNNs), due to their powerful ability to extract spatial features from image and video data, are highly effective for human activity recognition from visual content. Particularly for activities involving complex motions, employing CNNs can significantly enhance recognition accuracy. In the proposed method, the filter size, number of filters, and stride length were determined using the Particle Swarm Optimization (PSO) algorithm. Simulation results indicate the high efficiency and performance of the proposed method.

Declaration

We acknowledge that we used ChatGPT to enhance the academic writing of our manuscript while ensuring the originality and integrity of our work.

Transparency Statement

The data supporting this study are available upon reasonable request to the corresponding author, subject to ethical and confidentiality considerations.

Acknowledgments

We would like to express our gratitude to all individuals who contributed to this project.

Declaration of Interest

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

REFERENCES

[1] Shajari, S., Kuruvinashetti, K., Komeili, A., & Sundararaj, U. (2023). The emergence of AI-based wearable sensors for digital health technology: a review. Sensors, 23(23), 9498.https://doi.org/10.3390/s23239498

[2] Alzubaidi, L., Zhang, J., Humaidi, A., Al-Dujaili, A., Duan, Y., Al-Shamma, O., … Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8, Article 53. https://doi.org/10.1186/s40537-021-00444-8

[3] Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys, 46(3), Article 33, 33:1–33:33. https://doi.org/10.1145/2499621

[4] Schuldt, C., Laptev, I., & Caputo, B. (2004). Recognizing human actions: A local SVM approach. In Proceedings of the 17th International Conference on Pattern Recognition (ICPR) (pp. 32–36). IEEE. https://doi.org/10.1109/ICPR.2004.1334462

[5] Bott, O. J., Ammenwerth, E., Brigl, B., Knaup, P., Lang, E., Pilgram, R., & Kulikowski, C. (2005). The challenge of ubiquitous computing in health care: Technology, concepts and solutions. Methods of Information in Medicine, 44(3), 473–479. https://doi.org/10.1055/s-0038-1633996

[6] Kumar, P., Chauhan, S., & Awasthi, L. (2023). Human activity recognition (HAR) using deep learning: Review, methodologies, progress and future research directions. Archives of Computational Methods in Engineering, 31, 179–219. https://doi.org/10.1007/s11831-023-09986-x

[7] Ronao, C. A., & Cho, S.-B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 59, 235–244. https://doi.org/10.1016/j.eswa.2016.04.032

[8] Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors, 16(1), Article 115. https://doi.org/10.3390/s16010115

[9] Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. Pattern Recognition Letters, 119, 3–11. https://doi.org/10.1016/j.patrec.2018.02.010

[10] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. M. (2016). Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors, 16(4), Article 426. https://doi.org/10.3390/s16040426

[11] Hassan, M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. (2018). A robust human activity recognition system using smartphone sensors and deep learning. Future Generation Computer Systems, 81, 307–313. https://doi.org/10.1016/j.future.2017.11.029

[12] Chen, C., Jafari, R., & Kehtarnavaz, N. (2017). A survey of depth and inertial sensor fusion for human action recognition. Multimedia Tools and Applications, 76(3), 4405–4425. https://doi.org/10.1007/s11042-015-3177-1

[13] Ray, A., Kolekar, M. H., Balasubramanian, R., & Hafiane, A. (2023). Transfer learning enhanced vision-based human activity recognition: A decade-long analysis. International Journal of Information Management Data Insights, 3(1), Article 100142. https://doi.org/10.1016/j.jjimei.2022.100142

[14] Dawar, N., & Kehtarnavaz, N. (2018, June). A convolutional neural network-based sensor fusion system for monitoring transition movements in healthcare applications. In 2018 IEEE 14th International Conference on Control and Automation (ICCA) (pp. 482–485). IEEE. https://doi.org/10.1109/ICCA.2018.8444326

[15] Wang, H., Zhao, J., Li, J., Tian, L., Tu, P., Cao, T., … & Li, S. (2020). Wearable sensor-based human activity recognition using hybrid deep learning techniques. Security and Communication Networks, 2020, Article 2132138. https://doi.org/10.1155/2020/2132138

[16] Lara, O. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. IEEE Communications Surveys & Tutorials, 15(3), 1192–1209. https://doi.org/10.1109/SURV.2012.110112.00192

TandfonlineSpace Frontiers

[17] Xie, X.-F., Zhang, W.-J., & Yang, Z.-L. (2003). Overview of particle swarm optimization. Control and Decision, 18(2), 129–134.

[18] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)* , 265-283.