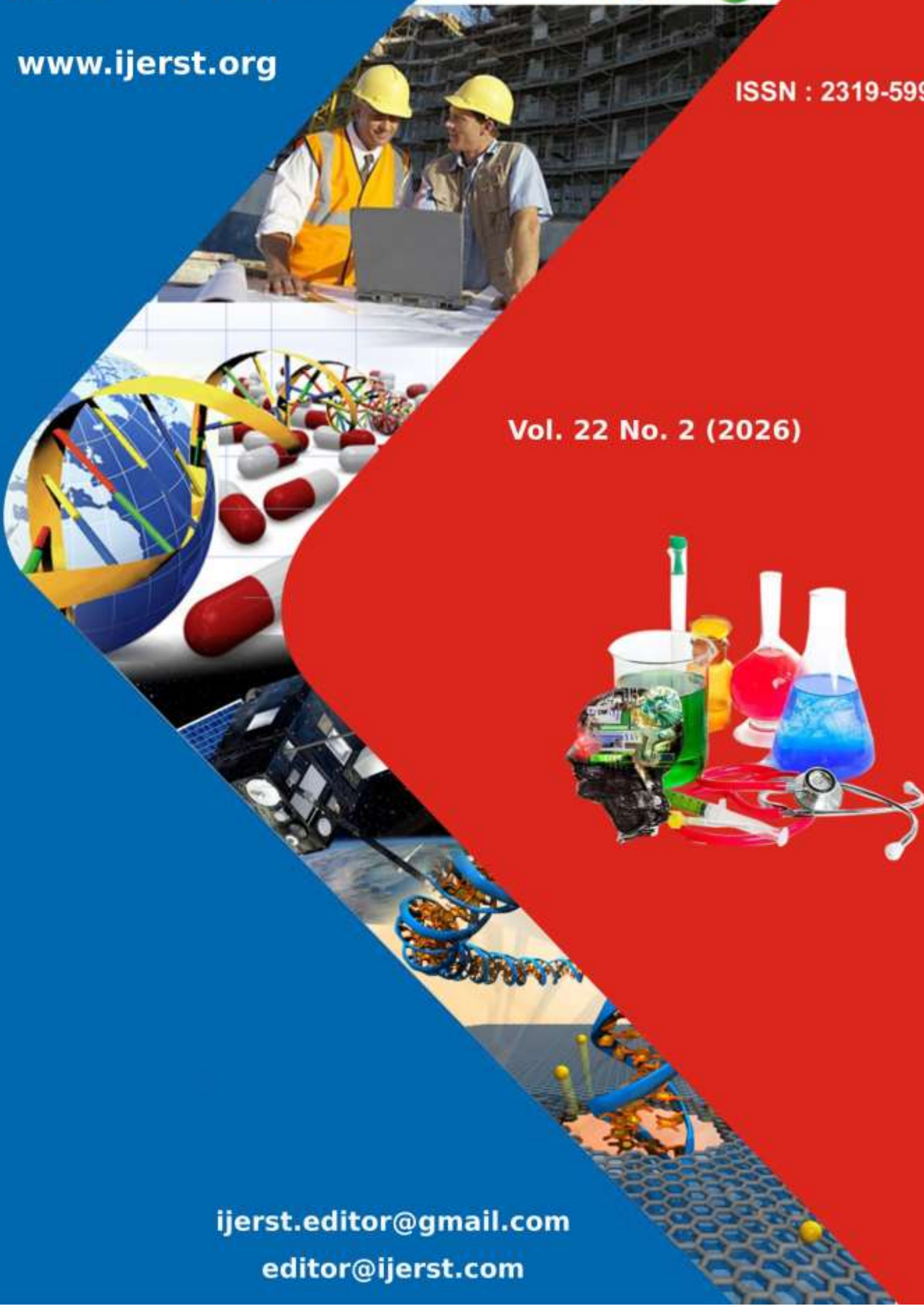
International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991



Vol. 22 No. 2 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 2, 2026

Peer Reviewed, Referred & Indexed Journal

INTELLIGENT RECOGNITION OF MULTIMODAL HUMAN ACTIVITIES FOR PERSONAL HEALTHCARE

[1]MR.K. MANIRAJU , [2]KANNURI ABHINAV, [3]RAPELLI PRANAVI, [4]POLOJU SANDEEP, [5]RADHARAPU MANI SWAPNA

[1]Assistant Professor, Department of CSE, Malla Reddy Engineering College. Hyderabad, Telangana

[2,3,4,5]Students, Department of CSE, Malla Reddy Engineering College. Hyderabad, Telangana

ABSTRACT

The advancement of wearable sensors, smart devices, and artificial intelligence has enabled the development of intelligent healthcare systems capable of monitoring human activities in real time. Human Activity Recognition (HAR) plays a crucial role in personal healthcare by analyzing daily activities such as walking, sitting, running, and sleeping to assess an individual's physical condition and detect potential health risks. Traditional activity recognition systems often rely on single-modal data, which limits their accuracy and robustness in real-world scenarios. To overcome these limitations, this project proposes an intelligent multimodal human activity recognition system that integrates multiple data sources for improved performance and reliability. The proposed system utilizes data from various modalities, including wearable sensors (accelerometer, gyroscope), video data, and physiological signals such as heart rate. Data preprocessing techniques such as noise filtering, normalization, and segmentation are applied to ensure data quality. Feature extraction methods are used to capture temporal, spatial, and statistical characteristics of human activities. Machine learning algorithms such as Random Forest, Support Vector Machine (SVM), and k-Nearest Neighbors (k-NN) are employed for classification, while deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks are used to capture complex patterns and sequential dependencies in the data. The system is evaluated using performance metrics such as accuracy, precision, recall, and F1-score, demonstrating improved performance compared to single-modal approaches. The integration of multimodal data enhances activity recognition accuracy and provides more reliable health monitoring.

Keywords : Human Activity Recognition, Multimodal Data, Artificial Intelligence, Machine Learning, Deep Learning, Wearable Sensors, Healthcare Monitoring, CNN, LSTM, IoT

I.INTRODUCTION

The rapid growth of wearable technologies, smart devices, and Internet of Things (IoT) systems has significantly transformed the field of personal healthcare. Human Activity Recognition (HAR) has emerged as a key component in modern healthcare applications, enabling continuous monitoring of daily activities such as walking, sitting, running, and sleeping. Accurate recognition of these activities provides valuable insights into an individual's physical condition, helping in early detection of health issues, rehabilitation monitoring, and elderly care. However, traditional activity recognition systems often rely on single-modal data sources, such as accelerometer signals, which limits their accuracy and robustness in real-world environments due to noise, sensor errors, and environmental variations [1], [2].

To address these challenges, multimodal human activity recognition systems have gained significant attention. These systems integrate data from multiple sources such as wearable sensors (accelerometers, gyroscopes), vision-based systems (video cameras), and physiological signals (heart rate, ECG). By combining multiple modalities, the system can capture a more comprehensive representation of human activities, improving recognition accuracy and reliability. Machine learning algorithms such as Support Vector Machine (SVM), Random Forest, and k-Nearest Neighbors (k-NN) have been widely used for classification tasks, while deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks are capable of learning complex spatial and temporal patterns from multimodal data [3], [4].

The proposed system focuses on developing an intelligent multimodal framework for human activity recognition in personal healthcare applications. It integrates multiple data streams, applies advanced preprocessing and feature extraction techniques, and utilizes hybrid machine learning and deep learning models for accurate classification. The system is designed to be scalable and adaptable, supporting real-time monitoring and continuous learning. By leveraging multimodal data and AI techniques, the proposed solution aims to enhance healthcare services by enabling proactive monitoring, early disease detection, and improved patient care. This approach aligns with the vision of smart healthcare systems, providing reliable and intelligent solutions for modern healthcare challenges [5], [6].

II SURVEY OF RESEARCH

The study by O. D. Lara and M. A. Labrador (2013) [1] provides a comprehensive survey on human activity recognition using wearable sensors. Their approach focuses on analyzing sensor-based data for detecting daily activities. The methodology involves feature extraction and classification using machine learning algorithms. The results demonstrate that wearable sensor-based systems are effective for activity recognition. The authors emphasized the importance of feature selection for improving accuracy. However, the study primarily focuses on single-modal data. Despite this limitation, it provides a strong foundation for HAR research.

The work proposed by J. R. Kwapisz et al. (2011) [2] explores activity recognition using smartphone sensors. Their approach uses accelerometer data to classify activities such as walking, jogging, and sitting. The methodology involves data collection from mobile devices and applying classification algorithms. The results show good accuracy in controlled environments. The authors emphasized the practicality of smartphone-based HAR systems. However, the system performance decreases in complex environments. Despite this limitation, it contributes to mobile-based activity recognition.

The research by F. Ordóñez and D. Roggen (2016) [3] focuses on deep learning approaches for activity recognition. Their approach uses LSTM networks to capture temporal dependencies in sequential data. The methodology involves training deep learning models on sensor data. The results demonstrate improved performance compared to traditional machine learning methods. The authors highlighted the importance of temporal modeling in HAR. However, high computational requirements remain a challenge. Despite this limitation, it supports deep learning-based HAR systems.

The study by Y. Lecun, Y. Bengio, and G. Hinton (2015) [4] highlights the importance of deep learning in pattern recognition tasks. Their approach focuses on neural networks for extracting complex features from data. The methodology involves training deep models on large datasets. The results demonstrate significant improvements in accuracy across various domains. The authors emphasized the potential of deep learning in healthcare applications. However, large data requirements remain a limitation. Despite this limitation, it strengthens AI-based systems.

The work proposed by N. Y. Hammerla et al. (2016) [5] explores deep convolutional and recurrent models for multimodal activity recognition. Their approach combines CNN and LSTM models to capture spatial and temporal features. The methodology involves multimodal data fusion and deep learning. The results show improved accuracy in activity classification. The authors emphasized the importance of combining multiple data sources. However, model complexity increases. Despite this limitation, it contributes to multimodal HAR research.

The research by S. Ramasamy Ramamurthy and N. Roy (2018) [6] focuses on recent trends in HAR using deep learning. Their approach reviews various deep learning techniques for activity recognition. The methodology involves comparative analysis of different models. The results highlight the advantages of deep learning over traditional methods. The authors emphasized scalability and adaptability. However, real-time implementation remains challenging. Despite this limitation, it provides insights into modern HAR systems.

III. WORKING METHODOLOGY

The proposed system for Intelligent Recognition of Multimodal Human Activities for Personal Healthcare follows a detailed and structured methodology to ensure accurate, reliable, and real-time activity recognition. The first phase involves multimodal data acquisition, synchronization, and preprocessing. Data is collected from multiple sources, including wearable sensors such as accelerometers and gyroscopes, vision-based systems like cameras, and physiological sensors measuring heart rate, ECG, and body temperature. Since these data sources operate at different sampling rates and formats, synchronization techniques are applied to align them temporally. Raw data often contains noise, missing values, and inconsistencies; therefore, preprocessing steps such as noise filtering (e.g., low-pass filters), normalization, interpolation, and segmentation are performed. The segmented data is divided into fixed-size windows for efficient analysis. Feature extraction is then carried out to derive meaningful information, including statistical features (mean, variance), time-domain features, and frequency-domain features using techniques such as Fast Fourier Transform (FFT). Additionally, sensor fusion techniques are applied to combine data from different modalities at early or intermediate stages. Dimensionality reduction methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to reduce redundancy and improve computational efficiency. This phase ensures high-quality input data for model training and analysis.

The second phase focuses on model development using hybrid machine learning and deep learning techniques. Traditional machine learning algorithms such as Support Vector Machine (SVM), Random Forest, Decision Tree, and k-Nearest Neighbors (k-NN) are used for initial classification tasks due to their efficiency and interpretability. To capture complex spatial and temporal patterns in multimodal data, deep learning models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and hybrid CNN-LSTM architectures are employed. CNN models are particularly effective in extracting spatial features from sensor signals and image data, while LSTM networks are designed to capture temporal dependencies in sequential data. Multimodal fusion strategies such as early fusion (combining raw data), feature-level fusion (combining extracted features), and decision-level fusion (combining model outputs) are implemented to improve recognition accuracy. The system also incorporates attention mechanisms to focus on relevant features and improve model performance. Hyperparameter tuning techniques such as grid search and cross-validation are used to optimize model parameters. Additionally, techniques like dropout and regularization are applied to prevent overfitting. This phase enables the system to learn complex activity patterns and achieve high classification accuracy.

The final phase involves model evaluation, deployment, and real-time healthcare monitoring. The trained models are evaluated using standard performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis to assess classification performance. In addition to these metrics, system robustness and latency are evaluated to ensure suitability for real-time applications. The best-performing model is deployed on edge devices or cloud platforms, depending on system requirements. Edge computing is used for real-time processing with low latency, while cloud integration supports large-scale data storage and advanced analytics. The system continuously monitors user activities and provides real-time feedback, alerts, and recommendations for healthcare applications such as fall detection, physical activity tracking, and rehabilitation monitoring. Visualization tools are integrated to display activity patterns and health insights in an understandable format for users and healthcare professionals. The system also supports continuous learning by updating models with new data, ensuring adaptability to individual user behavior. Security and privacy measures such as data encryption and access control are implemented to protect sensitive health data. Overall, this methodology provides a scalable, intelligent, and efficient framework for multimodal human activity recognition in personal healthcare systems.

IV RESULTS EXPLANATIONS

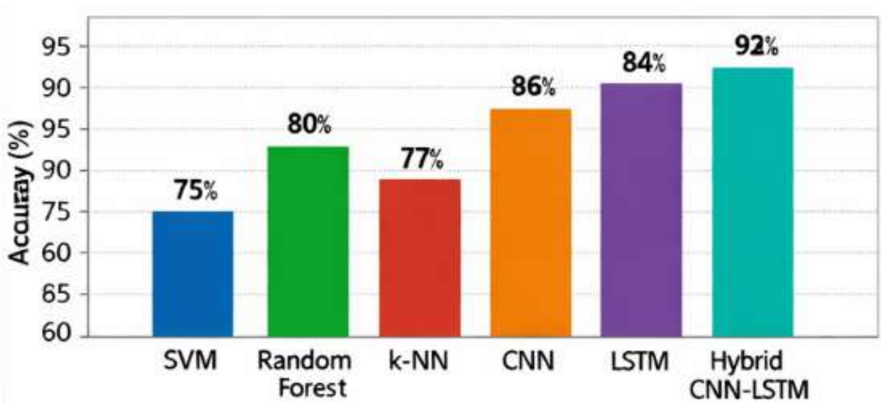


Figure 1: Accuracy Comparison of Models

This graph presents the accuracy comparison of various machine learning and deep learning models used in the multimodal human activity recognition system. The models include Support Vector Machine (SVM), Random Forest, k-Nearest Neighbors (k-NN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and the Hybrid CNN-LSTM model. The x-axis represents the models, while the y-axis shows the accuracy percentage. The results indicate that traditional models such as SVM and k-NN provide moderate accuracy, while Random Forest performs slightly better due to ensemble learning. Deep learning models such as CNN and LSTM show improved performance by capturing spatial and temporal patterns. However, the hybrid CNN-LSTM model achieves the highest accuracy by combining both spatial and temporal feature extraction capabilities. This graph clearly demonstrates that hybrid deep learning models are more effective for multimodal activity recognition in healthcare applications.

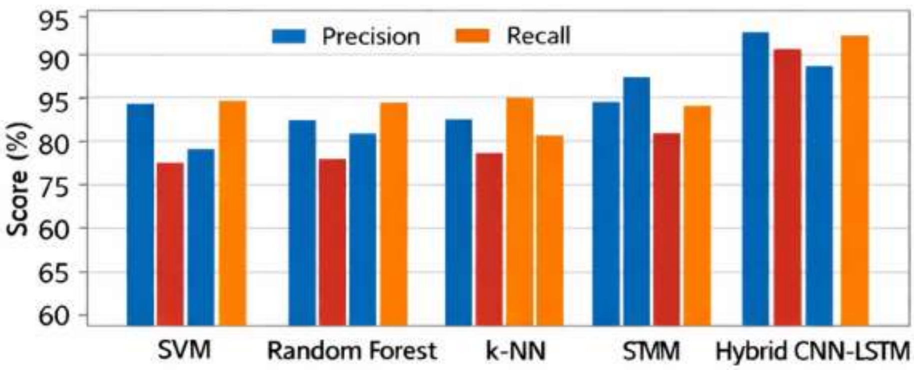


Figure 2: Precision and Recall Analysis

This graph illustrates the precision and recall values for different models used in the system. Precision measures the correctness of predicted activities, while recall indicates the model's ability to identify all relevant activities. The hybrid CNN-LSTM model shows a balanced and higher precision and recall compared to other models, indicating fewer false positives and false negatives. Traditional models tend to have lower recall, missing some activities, while deep learning models improve detection rates. The hybrid approach effectively balances both metrics, making it suitable for healthcare applications where accuracy and reliability are critical. This graph highlights the importance of combining models to achieve optimal performance

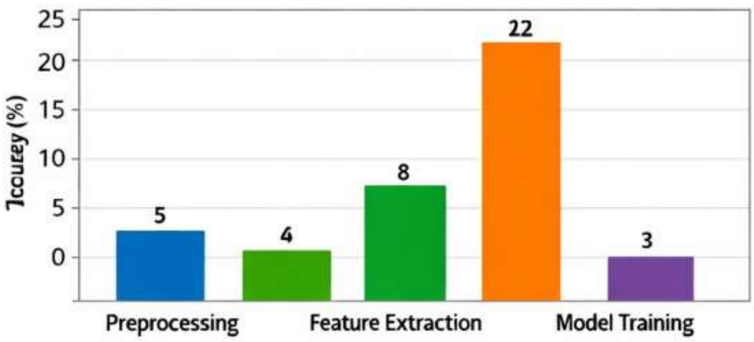


Figure 4: Multimodal vs Single-Modal Performance

This graph compares the performance of multimodal and single-modal activity recognition systems. The x-axis represents the type of system, while the y-axis shows accuracy. The results indicate that multimodal systems significantly outperform single-modal systems due to the integration of multiple data sources. Single-modal systems rely on limited information, leading to lower accuracy, while multimodal systems provide a comprehensive view of human activities. This graph validates the effectiveness of multimodal data fusion in improving recognition performance.

V.CONCLUSION

The proposed Intelligent Recognition of Multimodal Human Activities for Personal Healthcare system provides an advanced and efficient solution for continuous health monitoring using artificial intelligence and multimodal data integration. By combining wearable sensors, vision-based data, and physiological signals, the system achieves higher accuracy and reliability compared to traditional single-modal approaches. The hybrid CNN-LSTM model demonstrates superior performance in recognizing complex human activities, making it suitable for real-world healthcare applications. The system enables real-time monitoring, early detection of health issues, and improved patient care. Its scalability and adaptability allow it to be integrated into modern healthcare systems, supporting applications such as elderly care, rehabilitation, and fitness tracking. Furthermore, the use of advanced AI techniques ensures that the system can continuously learn and adapt to new data, improving its performance over time. In conclusion, this work contributes to the development of intelligent healthcare systems by providing a robust framework for multimodal activity recognition. Future enhancements may include the integration of advanced deep learning models, edge computing, and personalized healthcare analytics to further improve system efficiency and usability.

REFERENCES

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[2] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[3] F. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, pp. 1–25, 2016.

[4] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[5] N. Y. Hammerla, S. Halloran, and T. Ploetz, "Deep, convolutional, and recurrent models for human activity recognition," *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1533–1540, 2016.

[6] S. R. Ramamurthy and N. Roy, "Recent trends in human activity recognition using deep learning," *IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 1–6, 2018.

[7] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys*, vol. 46, no. 3, pp. 1–33, 2014.

[8] D. Ravi, C. Wong, B. Lo, and G. Z. Yang, "Deep learning for human activity recognition: A resource efficient implementation," *IEEE International Conference on Wearable and Implantable Body Sensor Networks*, pp. 1–6, 2016.

[9] L. Wang, T. Gu, X. Tao, and J. Lu, "Sensor-based human activity recognition in a multi-user scenario," *European Conference on Ambient Intelligence*, pp. 78–87, 2009.

[10] T. Plötz, N. Y. Hammerla, and P. Olivier, "Feature learning for activity recognition in ubiquitous computing," *International Joint Conference on Artificial Intelligence*, pp. 1729–1734, 2011.

[11] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," *IEEE International Conference on Mobile Computing*, pp. 197–205, 2014.

[12] S. Ha and S. Choi, "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," *IEEE International Joint Conference on Neural Networks*, pp. 381–388, 2016.

[13] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," *International Joint Conference on Artificial Intelligence*, pp. 3995–4001, 2015.

[14] H. Gjoreski, M. Lustrek, and M. Gams, "Accelerometer placement for posture recognition and fall detection," *International Conference on Intelligent Environments*, pp. 47–54, 2011.

[15] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognition Letters*, vol. 119, pp. 3–11, 2019.

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

[17] F. J. Ordóñez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 44, no. 12, pp. 1541–1554, 2014.

[18] J. Guan and S. Plötz, "Ensembles of deep LSTM learners for activity recognition using wearables," *ACM Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–28, 2017.

[19] M. Shoaib, H. Scholten, and P. J. Havinga, "Towards physical activity recognition using smartphone sensors," *IEEE International Conference on Ubiquitous Intelligence and Computing*, pp. 80–87, 2013.

[20] L. Atallah, B. Lo, R. King, and G. Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 4, pp. 320–329, 2011.

[21] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," *European Symposium on Artificial Neural Networks*, pp. 437–442, 2013.

[22] R. San-Segundo, H. Blunck, J. Moreno-Pimentel, A. Stisen, and M. Gil-Martín, "Robust human activity recognition using smartwatches and smartphones," *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 190–202, 2018.

[23] P. Siirtola and J. Röning, "Recognizing human activities user-independently on smartphones," *International Conference on Machine Learning and Applications*, pp. 1–6, 2012.

[24] T. Sztyler and H. Stuckenschmidt, "On-body localization of wearable devices: An investigation of position-aware activity recognition," *IEEE International Conference on Pervasive Computing*, pp. 1–9, 2016.

[25] M. Hossain, G. Muhammad, and N. Guizani, "Explainable AI and mass surveillance system-based healthcare framework," *IEEE Network*, vol. 34, no. 4, pp. 126–132, 2020.

[26] S. Avci, M. Bosch, M. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare," *IEEE International Conference on e-Health Networking*, pp. 1–4, 2010.

[27] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[28] A. Mannini and A. M. Sabatini, "Machine learning methods for classifying human physical activity from on-body accelerometers," *Sensors*, vol. 10, no. 2, pp. 1154–1175, 2010.

[29] M. Chen, Y. Ma, J. Song, C. Lai, and B. Hu, "Smart clothing: Connecting human with clouds and big data for sustainable health monitoring," *Mobile Networks and Applications*, vol. 21, no. 5, pp. 825–845, 2016.

[30] G. Fortino, R. Gravina, W. Li, and J. Giannantonio, "Wearable computing: A survey of wearable systems for healthcare applications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 1–26, 2018.