Journal for Educators, Teachers and Trainers
The PROFESIOLab electronic, peer-reviewed, open-access Magazine

ISSN 1989-9572

DOI:10.47750/jett.2024.15.05.33

Machine Learning-based Multi-Class Classification of Human Fitness Activities for Personalized Wellness Solutions

B. Vasantha1, Pandiri Sanjana2, Survi Nishitha Goud2, Thukkani Mounika2

Journal for Educators, Teachers and Trainers, Vol.15(5)

https://jett.labosfor.com/

Date of Reception: 24 Oct 2024

Date of Revision: 20 Nov 2024

Date of Publication : 31 Dec 2024

B. Vasantha1, Pandiri Sanjana2, Survi Nishitha Goud2, Thukkani Mounika2 (2024). Machine Learning-based Multi-Class Classification of Human Fitness Activities for Personalized Wellness Solutions,Vol.15(5).338-345

Journal for Educators, Teachers and Trainers
The LabOSfor electronic, peer-reviewed, open-access Magazine

Journal for Educators, Teachers and Trainers, Vol. 15(5)

ISSN1989 –9572

https://jett.labosfor.com/

Machine Learning-based Multi-Class Classification of Human Fitness Activities for Personalized Wellness Solutions

B. Vasantha[1], Pandiri Sanjana[2], Survi Nishitha Goud[2], Thukkani Mounika[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Information Technology

[1,2]Malla Reddy Engineering College for Women (UGC – Autonomous), Maisammaguda, Hyderabad, 500100, Telangana.

Corresponding Email: vasanthait63@gmail.com

Abstract

The growth of sedentary lifestyles and lifestyle-related health conditions has highlighted the need for individualized wellness solutions to promote physical activity and health. Basic exercises like jumping jacks and squats to complex routines like pull-ups promote physical health and personalized wellness. General advice or manual tracking methods could not provide enough precision and customisation to meet individual fitness needs. Manual fitness diaries, basic workout regimens, and standardized fitness exams gave limited user performance and progress insights, sometimes resulting in unsatisfactory results. These labour-intensive, error-prone, and unadaptable methods made it hard to achieve goals or discover areas for improvement. This research uses new computational methods to improve fitness monitoring systems, inspired by the transformational potential of AI and ML in healthcare and fitness. Machine learning can create intelligent algorithms that classify human fitness activities using wearable technologies, sensor data, and massive datasets. Personal feedback, performance evaluation, and adaptable exercise plans are possible with this technique. Growing need for data-driven wellness solutions, ML algorithm breakthroughs, and scalable systems for different populations inspire us. Traditional systems struggle to interpret massive amounts of real-time data, recognize actions accurately, and scale personalized recommendations. Our suggested system classifies fitness activities using labeled sensor data and machine learning methods. This method provides accurate activity recognition and rich user performance insights. This comprehensive, automated, and scalable technology improves fitness routines, individualized wellbeing, and healthier lifestyles by solving existing approaches' shortcomings. This novel system could revolutionize fitness tracking and help achieve long-term wellness goals.

Keywords: Human lifestyle, Fitness solutions, Personalized wellness, Machine learning, Predictive analytics.

1. INTRODUCTION

In India, the prevalence of sedentary lifestyles has led to alarming rates of obesity, diabetes, and cardiovascular diseases, with over 135 million Indians affected by obesity as per recent studies. Despite increasing awareness about fitness, a significant portion of the population struggles with consistent physical activity due to lack of personalized fitness solutions. Historically, fitness monitoring was limited to subjective self-assessments or trainer evaluations, which lacked scalability and precision. With India being the second-largest consumer of wearable devices globally, there is a growing opportunity to harness data from these devices to provide customized wellness solutions. Combining technology and fitness can revolutionize health outcomes and empower individuals to achieve their wellness goals effectively. Machine learning-based systems for classifying human fitness activities enable precision, automation, and personalization in fitness monitoring. Applications range from activity recognition in wearable devices to tailored wellness plans for diverse users. This approach enhances healthcare, fitness tracking, and rehabilitation support, contributing to a healthier society.

Before the advent of machine learning, fitness activity classification relied on manual observation, subjective evaluations, and generic recommendations. These methods were time-consuming, prone to inaccuracies, and unable to scale for large populations. Personalized feedback was limited, leading to inefficient workout plans that often failed to meet individual needs.

The increasing adoption of wearable devices and advancements in sensor technology provide vast datasets for analyzing human activity. However, the challenge lies in extracting meaningful insights from this data. The motivation for this research stems from the potential to bridge this gap by using machine learning to automate activity classification, provide personalized feedback, and promote healthier lifestyles.

Traditional fitness monitoring systems relied on manual techniques such as handwritten logs, observational analysis by trainers, and generalized fitness plans. These methods lacked accuracy, scalability, and real-time feedback. The inability to adapt to individual needs and track performance efficiently led to suboptimal results and disengagement. The proposed system leverages machine learning algorithms, such as K-Nearest Neighbors (KNN) and Light Gradient Boosting Machine (LightGBM), to classify fitness activities using labeled sensor data. Research papers highlight the effectiveness of ML models in human activity recognition (HAR), with features like accelerometer and gyroscope data used for accurate classification. The system automates activity recognition, provides actionable feedback, and adapts to individual performance. Advances in deep learning, such as Convolutional Neural Networks (CNNs), can further enhance accuracy for more complex activities.

2. LITERATURE SURVEY

In [1] Breiman in his seminal work on "Random Forests" presents a highly efficient ensemble method for classification and regression problems, emphasizing its ability to handle high-dimensional datasets effectively. This approach leverages multiple decision trees to mitigate overfitting, enhance accuracy, and ensure robustness. The study underlines the versatility of Random Forests in diverse applications, setting a foundation for advanced machine learning algorithms. Rasmussen et al. [2] discussed the DELVE framework for validating learning experiments, offering a robust mechanism to evaluate the performance of machine learning models. Their work underscores the importance of a systematic approach in experimental machine learning, emphasizing reproducibility and empirical rigor in diverse data-driven studies. Alcalá-Fdez et al. [3] introduced the KEEL data-mining software, which integrates datasets, algorithms, and experimental frameworks for machine learning research. This tool simplifies the comparative analysis of algorithms, encouraging innovation in data preprocessing and classification tasks while fostering reproducibility in research.

In [4] Flake and Lawrence proposed an efficient Support Vector Machine (SVM) regression training methodology using Sequential Minimal Optimization (SMO). Their work highlights the computational advantages of SMO, making SVMs more accessible for large-scale applications with reduced training times. In [5] García and Herrera extended statistical comparison methodologies for classifiers, providing a framework for analyzing performance metrics across datasets. Their approach enhances the reliability of comparative studies, ensuring robust conclusions about algorithm efficiency and accuracy. Friedman presented a ranking-based statistical approach to address the normality assumptions inherent in traditional ANOVA [6]. His method remains foundational in evaluating multiple algorithms' performance without stringent statistical constraints, influencing contemporary research practices.

In [7] Kamel Boulos and Yang reviewed mobile applications for planning and tracking physical activities, emphasizing the need for integrated, user-friendly solutions. Their findings advocate for advancements in personalized activity tracking systems, addressing future challenges in mHealth technologies. Barshan and Yurtman [8] proposed a classification methodology for daily and sports activities using wearable sensors. Their work demonstrates robust activity recognition, invariant to sensor positioning, highlighting the potential of motion-sensing technologies in real-world applications. Yurtman et al. [9] extended activity recognition using quaternion-based transformations to mitigate wearable sensor orientation issues. Their study enhances the reliability of sensor-based systems, paving the way for improved accuracy in human activity tracking. In [10] Camomilla et al. provided a systematic review of wearable inertial sensors in sports performance evaluation. Their research identifies trends and challenges in real-time performance monitoring, advocating for advanced sensor technologies to bridge the gap between laboratory and field-based applications. Weygers et al. [11] presented a methodological review of inertial sensor-based kinematics for lower limb joint monitoring. Their findings emphasize the potential of sensor-based methodologies in biomechanical analysis, supporting advancements in rehabilitation and sports science. In [12] Jaouedi et al. proposed a hybrid deep learning model for human action recognition, combining convolutional and recurrent architectures. Their study underscores the advantages of deep learning in capturing temporal dependencies, enhancing the accuracy and generalizability of action recognition systems. Sun et al. [13] reviewed human action recognition using multiple data modalities, providing insights into challenges and advancements in fusing depth and inertial sensor data. Their comprehensive analysis outlines future directions for multi-modal action recognition systems, emphasizing the need for robust feature extraction and integration techniques.

4. PROPOSED SYSTEM

The proposed system aims to leverage machine learning algorithms to classify human fitness activities more accurately and efficiently. By comparing the performance of a traditional algorithm like K-Nearest Neighbors (KNN) with a state-of-the-art algorithm like Light Gradient Boosting Machine (LightGBM), this system seeks to address the limitations of traditional fitness tracking methods. The process involves utilizing labeled sensor data, preprocessing it for quality improvement, and applying classification algorithms to build a robust model. In this outlines the detailed steps, from dataset acquisition to prediction, and highlights the advantages of the proposed algorithm over existing ones.

Step 1: Dataset

The Dataset containing labelled sensor data collected from wearable devices. Each record in the dataset corresponds to specific fitness activities such as jumping jacks, push-ups, or squats, along with sensor readings like accelerometer and gyroscope values. This raw data serves as the input for building and testing machine learning models.

Step 2: Dataset Preprocessing

The raw dataset often contains noise, missing values, or irrelevant features. Preprocessing involves handling null values by removing or imputing them, normalizing data for consistency, and extracting relevant features. Label encoding is used to transform categorical labels into numerical format, enabling machine learning algorithms to interpret them.

Step 3: Label Encoder

The labels, which represent fitness activities, are converted into numerical form using a label encoder. For instance, activities like "push-up" or "jumping jacks" are encoded as integers. This ensures the data is machine-readable while preserving the relationship between labels.

Step 4: Existing Algorithm

The KNN algorithm is applied to the preprocessed dataset as a baseline classifier. KNN uses proximity-based calculations to classify activities by identifying the 'k' nearest data points in the feature space. Its simplicity makes it a good starting point, but its performance diminishes with large datasets and noisy data.

Step 5: Proposed Algorithm

LightGBM is implemented as the advanced classifier due to its ability to handle large-scale data efficiently. Unlike KNN, it constructs decision trees iteratively, optimizing for both speed and accuracy. LightGBM processes the dataset to identify patterns and classify activities with high precision.

Step 6: Performance Comparison

The models are evaluated using metrics like accuracy, precision, recall, and F1-score. A confusion matrix is used to visualize the classification performance, highlighting the strengths and weaknesses of each algorithm.

Step 7: Prediction

The LightGBM model is deployed to predict activity labels on unseen test data. The predictions are compared with true labels to assess real-world applicability.

4. RESULTS

Figure 1: Pre-processed Dataset Count Plot.

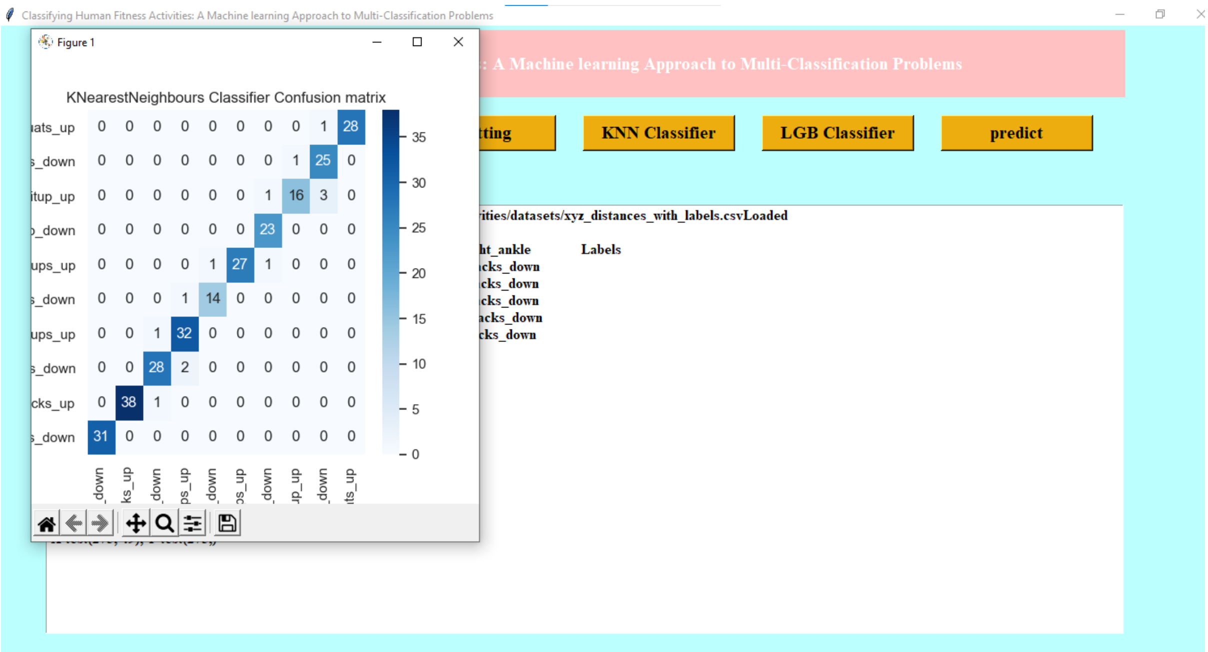

Figure 2: KNN Classifier Confusion matrix

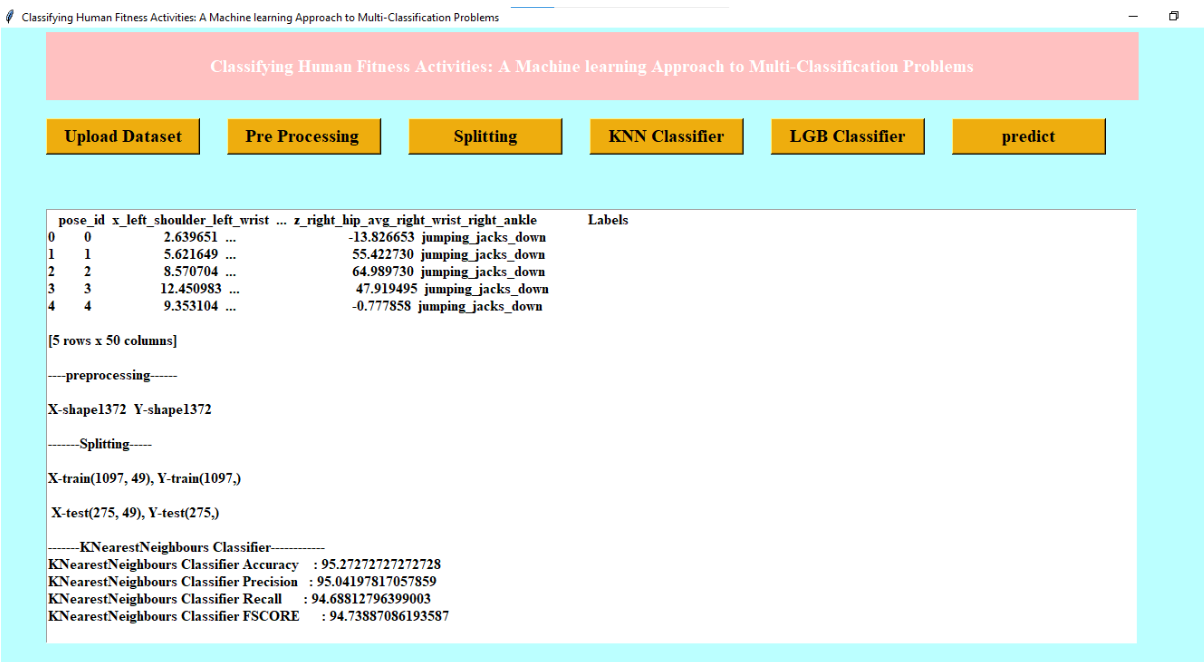

Figure 3: KNN Classifier Perfomance matrix

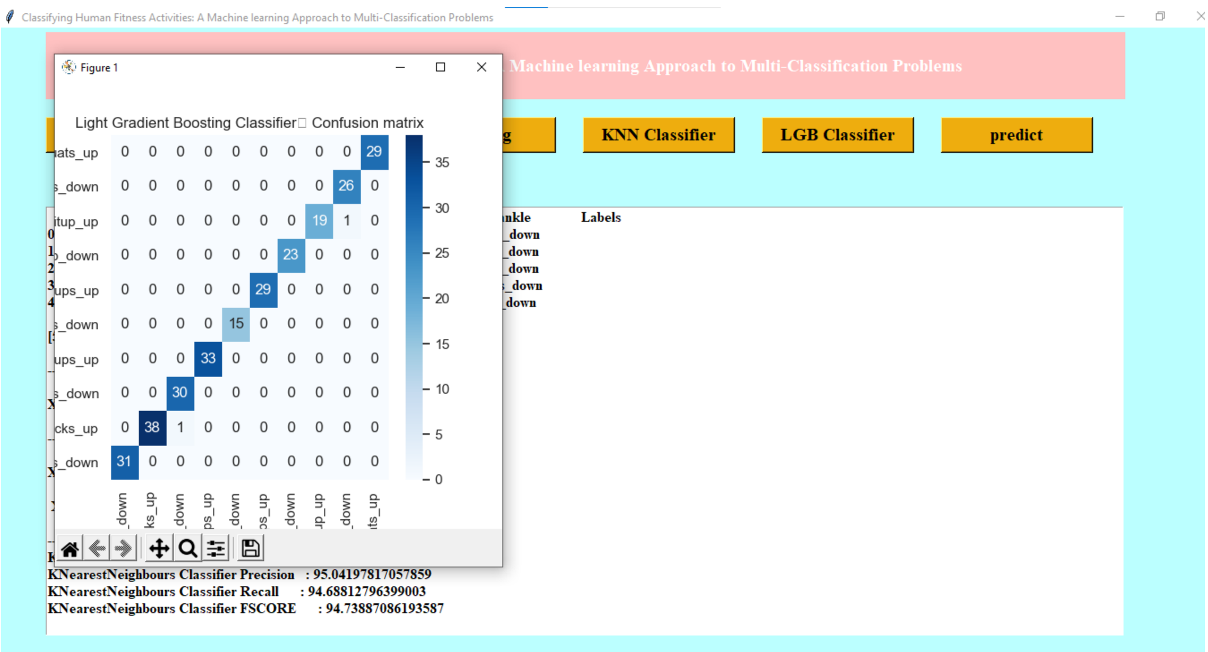

Figure 4: LightGB Classifier Confusion matrix.

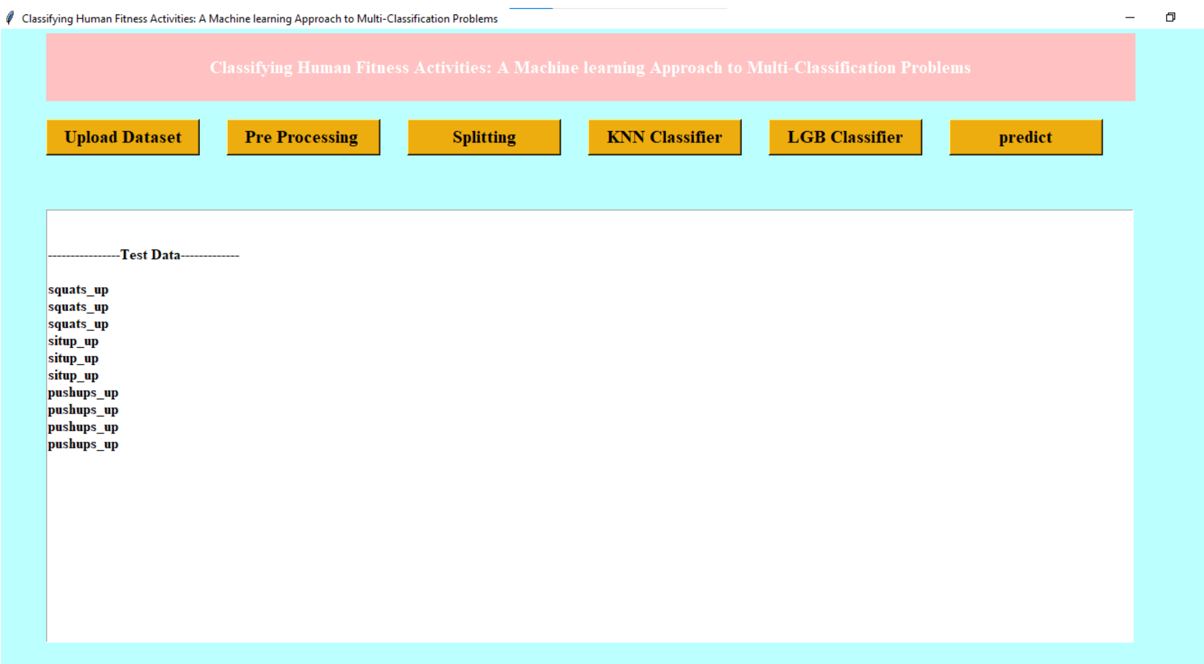

Figure 5: Predicted Output

5. CONCLUSIONS

The study and implementation of a machine-learning-based system for multi-class classification of human fitness activities highlight the importance of advanced computational methods in personalized wellness solutions. By leveraging labeled fitness activity data, the system successfully demonstrated the capability to classify a variety of activities using both K-Nearest Neighbors (KNN) and Light Gradient Boosting Machine (LightGBM) algorithms. The comparative analysis revealed the superior performance of LightGBM due to its ability to handle large datasets, manage categorical variables effectively, and provide faster training with high accuracy. This system bridges the gap between traditional manual monitoring methods and automated, data-driven solutions, offering a robust and scalable approach to fitness activity classification. The integration of data preprocessing, such as null value handling, label encoding, and scaling, ensures high-quality input for the machine learning models. The process of splitting the dataset into training and testing sets further strengthens the system's reliability, allowing for unbiased performance evaluation. The implementation provides a complete pipeline from data ingestion to activity prediction, which is crucial for real-time applications in fitness tracking devices, health monitoring systems, and wellness platforms. The results demonstrate that such a system can be pivotal in enhancing individual fitness routines, improving health outcomes, and promoting overall wellness. In conclusion, the paper establishes a solid foundation for future innovations in fitness activity classification, encouraging further exploration of machine learning in personalized healthcare.

REFERENCE

[1] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.K. Bache and M. Lichman, “UCI machine learning repository,” 2013.

[2] C. E. Rasmussen, R. M. Neal, G. Hinton, D. van Camp, M. Revow, Z. Ghahramani, R. Kustra, and R. Tibshirani, “Delve data for evaluating learning in valid experiments

[3] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, and S. García, “Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework,” Multiple-Valued Logic and Soft Computing, vol. 17, no. 2 3, pp. 255–287, 2011.

[4] G. W. Flake and S. Lawrence, “Efficient svm regression training with smo,” Machine Learning, vol. 46, no. 1-3, pp. 271–290, 2002.

[5] S. García and F. Herrera, “An extension on statistical comparisons of classifiers over multiple data sets for all pairwise comparisons,” Journal of Machine Learnin Research, vol. 9, no. 2677-2694, p. 66, 2008.

[6] M. Friedman, “The use of ranks to avoid the assumption of normality implicit in the analysis of variance,” Journal of American Statistical Association, vol. 32, pp. 675–701,1937

[7] M. N. Kamel Boulos and S. P. Yang, ‘‘Mobile physical activity planning and tracking: A brief overview of current options and desiderata for future solutions,’’ mHealth, vol. 7, p. 13, Jan. 2021, doi: 10.21037/mhealth.2020.01.01.

[8] B. Barshan and A. Yurtman, ‘‘Classifying daily and sports activities invariantly to the positioning of wearable motion sensor units,’’ IEEE Internet Things J., vol. 7, no. 6, pp. 4801–4815, Jun. 2020, doi: 10.1109/JIOT.2020.2969840.

[9] A. Yurtman, B. Barshan, and B. Fidan, ‘‘Activity recognition invariant to wearable sensor unit orientation using differential rotational transformations represented by quaternions,’’ Sensors, vol. 18, no. 8, p. 2725, Aug. 2018, doi: 10.3390/s18082725.

[10] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, ‘‘Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review,’’ Sensors, vol. 18, no. 3, p. 873, Mar. 2018, doi: 10.3390/s18030873.

[11] I. Weygers, M. Kok, M. Konings, H. Hallez, H. De Vroey, and K. Claeys, ‘‘Inertial sensor-based lower limb joint kinematics: A methodological systematic review,’’ Sensors, vol. 20, no. 3, p. 673, Jan. 2020, doi: 10.3390/s20030673. [173] C. Chen, R. Jafari, and N. Kehtarnavaz, ‘‘A survey of depth and inertial sensor fusion for human action recognition,’’ Multimedia Tools Appl., vol. 76, no. 3, pp. 4405–4425, Feb. 2017, doi: 10.1007/s11042-015-3177- 1.

[12] N. Jaouedi, N. Boujnah, and M. S. Bouhlel, ‘‘A new hybrid deep learning model for human action recognition,’’ J. King Saud Univ. Comput. Inf. Sci., vol. 32, no. 4, pp. 447–453, May 2020, doi: 10.1016/j.jksuci.2019.09.004.

[13] Z. Sun, Q. Ke, H. Rahmani, M. Bennamoun, G. Wang, and J. Liu, ‘‘Human action recognition from various data modalities: A review,’’ IEEE Trans. Pattern Anal. Mach. Intell., pp. 1–20, 2022, doi: 10.1109/TPAMI.2022.3183112.