IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 4, April 2026

Impact Factor: 8.2

Real-Time Fall and Activity Detection for Patient and Elderly Safety Using OpenCV

Prof. Dr. Mrunal Pathak, Satyam Bhagwat, Vijay Hissal, Yash Inamdar

Department of Information Technology

AISSMS Institute of Information Technology, Pune, Maharashtra, India

satyambhagwat30@gmail.com,vijayhissal3487@gmail.com, yashinamdar3349@mail.com

Abstract: *Vision-Based Human Activity Recognition (HAR) is an important area of computer vision that enables automatic detection and analysis of human actions such as walking, sitting, running, and falling using image and video data. It plays a vital role in smart healthcare systems, especially for elderly monitoring, fall detection, and assisted living support. With recent advancements in computer vision and deep learning, camera-based HAR methods have become more accurate and practical compared to traditional sensor-based approaches. This study presents a survey and framework overview of vision-based HAR using OpenCV integrated with machine learning and deep learning models. OpenCV provides efficient tools for video preprocessing, motion detection, silhouette extraction, and feature extraction, which support reliable real-time activity recognition. The work reviews existing HAR methodologies and applications in healthcare, smart homes, and surveillance, and identifies key challenges such as computational cost, privacy concerns, occlusion, lighting variation, and real-time performance constraints. The study emphasizes lightweight and privacy-aware vision pipelines and outlines future directions including deep learning optimization and edge deployment. The proposed approach aims to support efficient, real-time, and healthcare- focused activity recognition systems that improve safety and well-being.*

Keywords: Human Activity Recognition (HAR), Computer Vision, Smart Healthcare, Fall Detection, Vision-Based Monitoring

I. INTRODUCTION

Human Activity Recognition (HAR) has emerged as an important research area that integrates computer vision, artificial intelligence, and ubiquitous computing systems. The primary objective of HAR is to develop computational models capable of recognizing and classifying human activities using data obtained from images, video streams, and motion sensors. With the growing demand for automated and continuously monitored environments, HAR has gained significant importance across several real-world domains. These domains include healthcare systems for continuous patient monitoring and fall detection, intelligent surveillance systems for analyzing human behavior, and smart home environments that adapt device behavior based on user activity patterns. Beyond simple event detection, HAR focuses on the continuous and reliable under- standing of human activities in real time, thereby improving safety, efficiency, and the overall quality of human–machine interaction.

The development of effective HAR solutions, particularly for deployment on edge devices, depends heavily on efficient computer vision tools. One of the most widely used libraries in this field is OpenCV, which provides a comprehensive framework for image and video processing. OpenCV supports several fundamental operations such as video preprocessing, foreground extraction, motion tracking, and feature computation. These operations enable the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, April 2026
Impact Factor: 8.2

transformation of raw video data into structured features that can be further utilized by machine learning models for activity recognition.

Over time, the methodologies used for HAR have evolved significantly. Early approaches primarily relied on manually engineered features, including optical flow patterns, body silhouettes, and geometric motion descriptors. While these techniques performed reasonably well in controlled environments, they often struggled under real-world conditions such as variations in lighting, occlusions, and different camera viewpoints. To address these limitations, modern HAR systems employ more robust feature extraction techniques, motion analysis methods, and temporal modeling approaches that capture both spatial and temporal characteristics of human actions.

Despite these advancements, implementing HAR systems in real-world environments remains challenging. Variations in human behavior for similar actions can reduce classification consistency and affect system reliability. In addition, many accurate HAR models require significant computational resources, which makes deployment on resource-constrained devices difficult. Another important concern is model transparency, particularly for HAR systems used in sensitive applications such as healthcare and security. Therefore, continued research is required to improve the efficiency, transparency, and adaptability of HAR systems. This paper presents a review of recent developments in Human Activity Recognition, highlights the role of OpenCV in designing HAR systems, and discusses existing limitations along with potential future research directions for building reliable and efficient activity recognition systems.

II. LITERATURE REVIEW

A. Early Vision-Based Approaches

Early HAR models relied heavily on low-level feature extraction techniques such as motion history images, optical flow, and silhouette-based tracking. Traditional approachesfrequently used background subtraction and contour-based methods to differentiate between various human postures and movements. While these techniques were effective in con- trolled environments, their performance decreased significantly under dynamic lighting conditions, background clutter, and occlusions [6].

B. Machine Learning-Based HAR Models

With the emergence of machine learning techniques, several classification algorithms were introduced for activity recognition. Methods such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Hidden Markov Models (HMM) were widely used to classify extracted motion features. These approaches demonstrated improved generalization capabilities and relatively lower computational complexity compared to deep learning models. However, they still faced challenges when handling large-scale datasets, complex activity patterns, and highly dynamic environments[1].

C. Deep Learning and Hybrid Models

Recent developments in deep learning have significantly advanced the field of HAR. Convolutional Neural Networks (CNNs) are commonly used to extract spatial features from video frames, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are employed to capture temporal dependencies across sequential frames. Several researchers have also integrated OpenCV-based pre- processing techniques with CNN and LSTM models to im- prove recognition accuracy. These hybrid systems enable real- time activity recognition and demonstrate better robustness to variations in camera viewpoints and environmental conditions [8].

D. Role of OpenCV in Modern HAR

OpenCV plays an essential role as a preprocessing and visualization component in modern HAR systems. It supports real-time video acquisition, segmentation, and motion analysis through built-in APIs such as BackgroundSubtractorMOG2, Hough Transform, and pose estimation tools. Researchers frequently utilize OpenCV to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, April 2026
Impact Factor: 8.2

extract relevant features from video frames before passing the processed data to machine learning or deep learning models. This integration improves recognition accuracy and enables practical deployment of HAR systems in applications such as healthcare monitoring and intelligent surveillance.

Traditional Approaches to Industry Classification

III. SYSTEM ARCHITECTURE

Proposed System: The proposed Human Activity Recognition (HAR) system using OpenCV is designed with several integrated modules, each responsible for a specific function in the activity detection and classification process. The following are the key core modules of the system:

A. Video Acquisition

This captures real-time video using a webcam or CCTV camera. The captured video serves as the primary input for the activity recognition system.

B. Preprocessing Module

Video frames are processed using techniques such as grayscale conversion, noise reduction, and frame resizing. These steps help normalize lighting conditions and improve feature extraction.

C. Feature Extraction Module

Important visual and motion-based features are extracted from frames using OpenCV operations such as edge detection, contour detection, and motion analysis.

D. Activity Classification Module

Machine learning or deep learning models classify activities such as walking, sitting, standing, or falling based on extracted features.

E. Data Storage and Interface Module

Recognized activities are stored and displayed through a user interface that allows monitoring of activity logs and alerts

F. User Interface Module

Provides a graphical interface for caregivers or healthcare staff to monitor live video feeds and activity alerts. Offers options to review past logs, adjust system settings, and manage users.

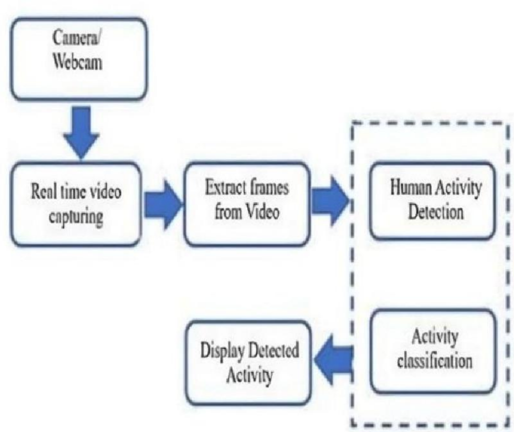

Fig.1 Process Flow for Human Activity Recognition

DOI: 10.48175/IJARSCT-32609

ISSN: 2581-9429 Volume 6, Issue 4, April 2026 Impact Factor: 8.2

The deployment diagram represents the physical deployment of the Human Activity Recognition system. This diagram can also help visualize the distribution of software components over hardware and thus help understand the operation of the system in a real-world environment. Major nodes of the deployment diagram are the Camera module for video acquisition, the Edge Processing Unit/Server for processing, and storage for storing the data. Caregivers and administrators can use any device, such as a computer, tablet, or smartphone, for accessing this system.

Background

Human Activity Recognition (HAR) is recognized as an essential research topic in the fields of computer vision and artificial intelligence, as it allows computers to recognize human activities. The increasing relevance of HAR stems from the need in various fields, including medical monitoring, security, and intelligent interactive systems. Essentially, the main objective of HAR is to recognize human activities through vision streams, thus enabling computers to perform appropriate actions in response. This segment discusses the concept of HAR, the system pipeline, and the role of OpenCV in developing efficient HAR systems.

Human Activity Recognition Overview

HAR is an interdisciplinary domain, combining the fields of image analysis, motion understanding, and learning-based prediction models, which can be used for the identification of physical activities such as walking, sitting, running, standing, and falling. It has been found that there are three basic stages in the conventional HAR process. These include data acquisition, feature extraction, and prediction. Data can be obtained through vision-based sensors and wearable devices. It is initially preprocessed to eliminate noise. After preprocessing, motion and posture indicators are computed through techniques such as shape boundary, pose, and motion vector. These are then converted into structured data, which can be understood by learning models. Learning models are used to predict activities based on learned patterns.

Importance of OpenCV in HAR

OpenCV is a highly popular and widely used open-source computer vision library that provides support for efficient and fast image and video processing operations. It provides optimized functions for frame transformations, foreground extraction, highlighting of motions, boundary detection, and object localization. In HAR systems, OpenCV is mainly utilized during preprocessing and feature preparation stages of the pipeline. It provides support for continuous frame processing, tracking of moving objects, and focused region extraction that contains relevant information regarding object motions. These operations help in eliminating redundant information from the background and improve the reliability of object classification. OpenCV can be integrated with modern learning libraries as well, enabling developers to create hybrid pipelines for object processing and neural inference.

Role of Machine Learning and Deep Learning

In the past, HAR systems implemented with traditional machine learning models based on handcrafted features were common. Decision Trees and ensemble models were popular choices when the feature vectors were explicitly engineered. However, with the recent advancements in HAR systems, deep learning models have gained popularity due to their ability to automatically learn the features directly from the input data. Convolutional Neural Networks can be effectively used to learn the spatial structure in images, whereas sequence- oriented models can be used to learn time-based dependencies in sequences of frames. Using OpenCV-based preprocessing, they can support stable real-time recognition in the presence of noise or variations in the environment.

Applications of HAR

Activity recognition technology can be applied in many real- life scenarios. In medical facilities, activity recognition can be used to study the movement patterns of patients and detect dangerous activities such as falling. In security systems, HAR can be used to study the activities of people and detect unusual activities. Smart home systems use

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, April 2026
Impact Factor: 8.2

activity understanding to control devices based on activities. Other applications of activity recognition technology include sports performance analysis, rehabilitation activity analysis, and industrial safety supervision. These scenarios show that HAR can be used in both personal and organizational settings.

Research Challenges in HAR

Although the performance of the HAR methods has been enhanced, some technical problems remain, which affect their performance. The performance of visual recognition may be influenced by factors such as differences in illumination, partial visibility, and differences in body shape and style. The requirement for real-time processing also presents some constraints, as the model has to perform efficiently on hardware with limited resources. Another challenge facing the field is the problem of ensuring the performance of the models in different environments and not just in a controlled environment.

IV. RESULT ANALYSIS

The evaluation of contemporary Human Activity Recognition systems indicates that this domain has significantly advanced in terms of prediction accuracy and processing speed. However, it is also evident that the performance of such systems is significantly affected by the learning models applied and the quality of features and preprocessing techniques involved. This is because different families of methods, such as rule-based and classical approaches, have different trade- offs in terms of performance and design considerations. This section will analyze these approaches in terms of performance and design considerations and how OpenCV-based processing is involved.

A. Comparison of HAR Methodologies

Methodology	Features Used	Tools / Algorithms	Accuracy	Real-Time	Limitations
Traditional / manual	Optical flow, contours, silhouettes	OpenCV background subtraction	Low	Yes	Unstable under lighting changes and occlusion
Machine learning based	Handcrafted feature vectors	SVM, k-NN, Decision Trees	Moderate	Partial	Dependent on feature quality and normalization
Deep learning based	Spatial features from frames	CNN, RNN, LSTM	~90%	Partial	High computational cost
Hybrid (CNN + LSTM)	Spatial + temporal features	CNN + LSTM + OpenCV	Low–mid (~90%)	Yes	Resource demanding on edge devices
OpenCV + deep learning	Preprocessed frame features	OpenCV + MobileNet	High	Yes	Complex scene handling difficult.

B. Contribution of OpenCV to System Efficiency

OpenCV is also useful as a practical enabling tool for HAR systems because it relieves some of the lower-level visual processing load. This is because it is highly optimized for fast video frame reading and transformation, segmentation, and motion-based filtering operations. This reduces some of the un- wanted visual information before it even reaches the classifier, making it easier on the model. Another benefit of OpenCV is its compatibility with newer learning frameworks. This makes it highly useful for creating a multilayered approach that lever- ages deterministic vision processing and neural networks for final recognition tasks. This is especially useful in situations where immediate responses are needed, such as in assisted living spaces and safety monitoring systems.

IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 6, Issue 4, April 2026
Impact Factor: 8.2

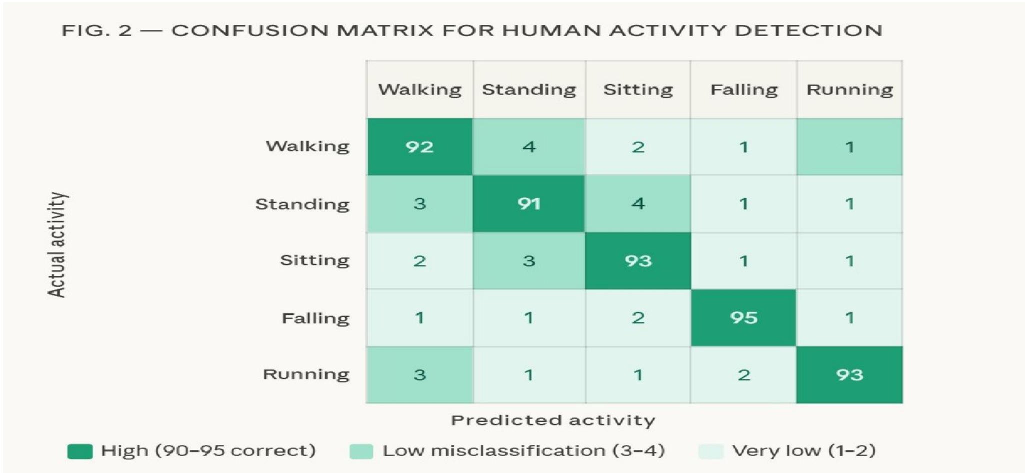

Fig.2 Confusion Matrix

C. Real-Time Processing Tradeoffs

InthecaseofoperationalHARsystems,itisfrequentlynec- essary for the systems to make predictions with the minimum delay. The lightweight nature of the vision pipelines, which mainlyinvolveclassicaloperationsbasedonOpenCV,makesit easy for the system to process the images with minimal delay, regardlessofthehardwareused.However,theprecisionofthe object recognition capability may be compromised for com- plex scenes. The predictive capabilities of deep neural models are usually superior, although the processing requirements for these models are more demanding. To address the trade-off between these requirements, developers have started to utilize optimized network families and hardware acceleration. The use of compact models like MobileNet and efficiency-focused networks minimizes the parameters used while maintainingthe required accuracy. Additional efficiency can be achieved by preprocessing the images to simplify the complexity of the images. The preprocessing steps include the extraction of the foreground and the cropping of the regions of interest.

D. Observed Limitations and Open Problems

However, despite the use of current technology, HAR systems are not exempt from shortcomings. In fact, the reliability of the recognition process can be compromised by factors such as shadows, clutter in the background, and the way different people perform the same task. The other problem with HAR systems is that they are prone to bias, especially because most of them are trained using limited and perhaps staged data sets and are not very good at generalizing toreal-world scenarios. The other problem with HAR systems is that they are prone to bias, especially because most of them are trained using limited and perhaps staged data sets and are not very good at generalizing to real-world scenarios. However, despite the use of current technology, HAR systems are not exempt from shortcomings. In fact, the reliability of the recognition process can be compromised by factors such as shadows, clutter in the background, and the way different people perform the same task. Improving the robustness of HAR systems in different environments will require improving the diversity of the training sets and using better augmentation.

E. Future Research Directions

Current analysis suggests that combining efficient vision preprocessing with adaptive deep learning remains a promising direction for next- generation HAR systems. One important path forward is multimodal recognition, where video cues are merged with wearable sensor signals to improve reliability. Edge and IoT-based processing architectures are also gaining importance because they reduce latency and bandwidth dependence. Another emerging focus area is explainable recognition models that can justify their predictions, which is particularly valuable in medical and security contexts. Continued work along these lines will support more transparent, scalable, and dependable activity recognition platforms.

V. CONCLUSION

Human Activity Recognition has evolved as a major en- abling technology for systems that must recognize human behavior automatically from visual or sensor-based informa- tion. The impact of HAR is evident in areas like medical supervision systems, smart security monitoring systems, and smart living spaces. The paper discussed how HAR techniques have moved from traditional manually crafted motion features to fully automated deep learning techniques that are able to recognize spatial and temporal patterns in human activity sequences. The paper also discussed how OpenCV provides a base for developing such systems through its support for fast video preprocessing and feature preparation. The integra- tion of OpenCV and learning frameworks allows for more effective design of accurate and deployable HAR systems. Especially, hybrid approaches combining convolution-based feature learning and temporal sequence learning have been observed to yield promising results for recognizing intricate actions from video streams. These hybrid approaches validate the idea of combining different modules of HAR for more robust and responsive HAR systems. However, at the same time, there are certain practical difficulties that still exist. For instance, the robustness of the model still depends on scene changes, visibility, and style variations among users. Another challenge associated with the model is the resource demand of the current advanced models, which makes it difficult for the model to be deployed at the edges. The future of the model will be influenced by the development of compact model structures, the integration of cross-modal learning with videos and wearable sensors, and the use of adaptive training forbetter generalization. In summary, HAR systems implemented using efficient vision libraries and intelligent learning models are moving towards wider deployment and application.

VI. FUTURE SCOPE

The field of Human Activity Recognition (HAR) is rapidly advancing, yet several opportunities remain for improvement and innovation. Future research should focus on enhancing system adaptability, accuracy, and real-time performance, particularly for healthcare and assistive technologies. The integration of OpenCV with emerging technologies such as deep learning, IoT, and edge computing will play a critical role in shaping next-generation HAR systems

A. Integration with IoT and Edge Devices

Deploying HAR on edge devices will reduce latency and enable real-time monitoring in healthcare facilities and homes.

B. Multi-modal Recognition Systems

Combining video with wearable sensors will improve recognition accuracy and enable smarter decision-making in healthcare applications.

C. Optimization for Low-Power Devices

Techniques like pruning and quantization will make HAR models efficient for deployment on resource-constrained and low-power devices.

REFERENCES

[1] J. Wang et al., “Deep Learning for Sensor-Based Human Activity Recognition,” Pattern Recognition Letters, 2019.
[2] S. Bhattacharya and P. Das, “Real-Time Human Activity Recognition Using OpenCV and CNN,” Journal of Ambient Intelligence, 2020.
[3] Krishna S. B. Siva et al., “Human Activity Recognition with OpenCV,” IJERT, 2024.
[4] I. A. Putra et al., “Human Action Recognition Using MediaPipe and LSTM,” Teknik Journal, 2022.
[5] K. Host and M. Ivasic-Kos, “Human Action Recognition in Sports Using Computer Vision,” Heliyon, 2022.

[6] M. Janapati et al., "Gait Driven Pose Tracking Using OpenCV," Engi- neering Proceedings, 2024.
[7] S. Gupta et al., "Human Action Recognition Using Deep Learning Techniques," Informatics and Automation, 2024.
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.
[9] M. Vrigkas, C. Nikou, and I. Kakadiaris, "A Review of Human Activity Recognition Methods," Frontiers in Robotics and AI, IEEE, 2015.
[10] L. Wang et al., "Temporal Segment Networks for Action Recognition in Videos," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.
[11] A. Jalal, S. Kamal, and D. Kim, "A Depth Video Sensor-Based Life-Logging Human Activity Recognition System for Elderly Care," Springer Journal of Sensors, 2014.
[12] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-Based Wearable Human Activity Recognition Using Misalignment-Robust Deep Learning," Springer Neural Computing and Applications, 2017.
[13] K. Simonyan and A. Zisserman, "Two-Stream Convolutional Networks for Action Recognition in Videos," NIPS, 2014.
[14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, MIT Press, 1997.
[15] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NeurIPS, 2012.