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 5, Issue 1, November 2025
Impact Factor: 7.67

Survey On: Human Activity Recognition using OpenCV

Prof. G. K. Gaikwad[1], Swami Bhagwat[2], Shreyash Vidhate[3], Darshan Shinde[4]
Professor, Department Artificial Intelligence and Data Science[1]
Student, Department Artificial Intelligence and Data Science[2,3,4]
MET Institute of Engineering, Nashik, Maharashtra, India

Abstract: *Human Activity Recognition (HAR) is an emerging field in computer vision that aims to automatically identify human actions such as walking, running, sitting, or falling using image and video analysis. This paper presents a comprehensive survey on HAR using OpenCV and machine learning techniques. OpenCV provides powerful tools for image preprocessing, motion detection, and feature extraction that enable accurate activity recognition. The study explores current methodologies, deep learning approaches, and key applications in healthcare, surveillance, and smart homes. It also discusses existing challenges such as occlusion, lighting variation, and real-time performance. The survey concludes by outlining potential research directions, including sensor fusion, deep learning integration, and edge deployment for enhanced HAR performance.*

Keywords*:* Human Activity Recognition (HAR), OpenCV, Machine Learning, Computer Vision, Real-Time Detection, Deep Learning

I. INTRODUCTION

Human Activity Recognition (HAR) is a rapidly evolving area within the fields of computer vision and artificial intelligence that focuses on detecting, analyzing, and classifying human actions through visual or sensor data. With the increasing demand for automation and intelligent monitoring systems, HAR has gained significant importance in various domains such as healthcare, surveillance, human-computer interaction, and smart home applications [1]. The primary goal of HAR is to enable computers to interpret human movements and behaviors in real-time, thereby improving safety, efficiency, and user experience. The advancement of computer vision frameworks, especially OpenCV, has made it possible to implement effective HAR systems even on low-cost hardware. OpenCV provides a wide range of tools for video analysis, object tracking, image preprocessing, and feature extraction, which are essential in identifying human motion patterns. These functionalities allow developers and researchers to design systems capable of distinguishing between activities like walking, sitting, standing, running, or falling, all of which hold immense potential in medical monitoring and assistive living environments. Traditional HAR systems relied heavily on handcrafted features such as optical flow, motion history images, and silhouette analysis. While these methods were effective to some extent, they were often limited by environmental factors like lighting conditions, occlusion, and background clutter. With the introduction of deep learning models and high-performance computing, the accuracy and robustness of HAR systems have improved dramatically. Integration of OpenCV with deep learning libraries such as TensorFlow and PyTorch has further enabled the development of hybrid models capable of extracting complex spatiotemporal features for real-time recognition. In healthcare applications, HAR has proven invaluable for continuous patient monitoring, fall detection among elderly individuals, and rehabilitation progress tracking [2]. In surveillance systems, HAR aids in detecting suspicious or abnormal activities, thereby enhancing security automation. Similarly, in smart home environments, HAR can be integrated with IoT systems to create context-aware automation solutions based on user behavior. Despite significant progress, several challenges remain in HAR implementation, such as variations in human posture, illumination changes, background dynamics, and computational constraints for real-time processing. This survey paper aims to analyze the evolution of HAR methodologies, highlight the role of OpenCV in this domain, review state-of-the-art approaches, and identify existing research gaps and potential directions for future work [4].

II. LITERATURE REVIEW INTRODUCTION

Human Activity Recognition (HAR) has been an active research domain for over a decade, evolving from simple motion detection to complex activity classification using deep learning models. Earlier approaches focused on traditional computer vision techniques involving handcrafted features, while modern systems employ deep neural networks capable of learning hierarchical motion representations. This section reviews key studies and methodologies related to HAR, highlighting their methods, datasets, strengths, and limitations.

A. Early Vision-Based Approaches

Initial HAR models relied heavily on low-level feature extraction methods such as motion history images, optical flow, and silhouette tracking. For example, traditional methods used background subtraction and contour-based analysis to differentiate between human postures. Although these approaches performed well under controlled environments, their accuracy degraded significantly under changing lighting conditions, occlusion, and complex backgrounds [6].

B. Machine Learning-Based HAR Models

With the introduction of machine learning, algorithms like Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Hidden Markov Models (HMM) were applied to classify extracted motion features. These systems showed better generalization capabilities and required less computational power compared to deep learning methods. However, they still struggled with dynamic background variations and large-scale datasets.

C. Deep Learning and Hybrid Models

Recent developments in deep learning have transformed HAR systems. Convolutional Neural Networks (CNNs) are used to extract spatial features from frames, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks model temporal dependencies between video frames. Researchers have also combined OpenCV-based preprocessing with CNN-LSTM architectures for enhanced accuracy. Such hybrid systems are capable of real-time recognition, robust against environmental changes, and adaptable to multiple camera perspectives.

D. Role of OpenCV in Modern HAR

OpenCV continues to play a crucial role as a preprocessing and visualization tool in HAR pipelines. It enables real-time video capture, frame segmentation, and skeletal tracking using modules such as BackgroundSubtractorMOG2, Hough Transform, and Pose Estimation APIs. Researchers commonly integrate OpenCV for feature extraction and feed the processed data into deep learning models, achieving high recognition rates in healthcare and surveillance applications.

III. MODULES AND BACKGROUND

Proposed Core Modules:

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:

Video Acquisition Module:

Captures live video input from a webcam or CCTV camera.

Ensure real-time video capture at a consistent frame rate.

Preprocessing Module:

Enhances the video frames by converting them to grayscale and reducing noise.

Normalizes lighting variations and resizes frames for consistent model input.

Feature Extraction Module:

Identifies important visual and motion-based features from frames using OpenCV techniques like edge detection,

contour extraction, and key point identification.
Prepares feature vectors that represent the activity in numerical form.

Activity Classification Module:
Applies a machine learning or deep learning model to classify activities such as walking, sitting, standing, or falling.
Integrates OpenCV's real-time analysis tools with trained models for immediate prediction.

Data Storage and Analysis Module:
Stores classified activity data and timestamps in a local or cloud-based database.
Supports visualization of patient activity trends and performance reports.

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.

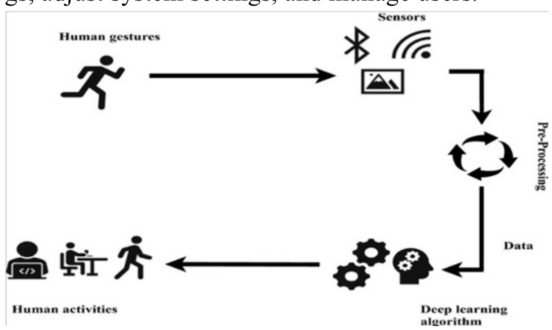

Figure 1. Process flow for HAR

The deployment diagram illustrates how the Human Activity Recognition sys- tem is physically deployed across hardware nodes and devices. It helps visualize the distribution of software components on hardware, showing how the system operates in a real-world environment. The main nodes include the Camera module for capturing video, an Edge Processing Unit or Server for preprocessing, feature extraction, and activity clas- sification, and storage units for logging activity data. Caregivers and administrators ac- cess the system through client devices such as computers, tablets, or smartphones. The diagram highlights communication channels, data flow, and deployment dependencies, ensuring proper system setup, scalability, and performance.

Background:
Human Activity Recognition (HAR) has become one of the most significant research areas in computer vision and artificial intelligence due to its vast range of applications in healthcare, security, and intelligent systems. The fundamental goal of HAR is to automatically identify human activities through video analysis or sensor data, enabling computers to understand and respond to human behavior effectively. This section presents an overview of the concepts, development, and importance of HAR, along with the role of OpenCV in building efficient recognition systems [6].

Human Activity Recognition Overview
Human Activity Recognition is a multidisciplinary field that combines image processing, motion analysis, and machine learning to classify various human actions such as walking, running, sitting, standing, or falling. The typical HAR system consists of three major phases: data acquisition, feature extraction, and activity classification. Data is collected through cameras or sensors, preprocessed to remove noise, and analyzed to detect motion patterns or postural changes. Feature extraction techniques such as contour analysis, skeleton modeling, and optical flow are used to identify key visual features that represent movement. Machine learning or deep learning algorithms then classify these features into specific activities based on training data [7].

Importance of OpenCV in HAR

OpenCV (Open Source Computer Vision Library) is an open-source framework widely used for image processing and video analysis. It provides a comprehensive set of functions for operations such as grayscale conversion, background subtraction, motion detection, edge detection, and contour analysis. In HAR systems, OpenCV plays a vital role in the early stages of video processing and feature extraction. It allows real-time frame analysis, object tracking, and region-of-interest detection, which form the foundation for accurate activity recognition. By integrating OpenCV with machine learning frameworks such as TensorFlow, Keras, or PyTorch, developers can design hybrid models that are capable of both traditional feature-based and deep learning-based recognition.

Role of Machine Learning and Deep Learning

Machine learning algorithms like Support Vector Machines (SVM), Random Forests, and Decision Trees have been widely used for classifying activities based on extracted features. However, in recent years, deep learning has significantly improved HAR accuracy by automatically learning hierarchical features from raw data. Models such as Convolutional Neural Networks (CNNs) capture spatial information from frames, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks capture temporal relationships between consecutive frames. When combined with OpenCV preprocessing, these models achieve robust real-time recognition even under challenging environmental conditions [8].

Applications of HAR:

The applications of Human Activity Recognition are vast and impactful. In healthcare, HAR systems monitor patients and detect abnormal activities like falls, reducing emergency response time. In surveillance, they identify suspicious movements or potential security threats. In smart homes, HAR enables automation based on resident behavior, improving comfort and safety. It is also used in sports analytics, rehabilitation tracking, and workplace safety monitoring. These applications demonstrate the adaptability and importance of HAR in both industrial and personal environments [3].

Research Challenges in HAR:

Despite advancements, HAR still faces challenges in achieving consistent accuracy across varied conditions. Factors such as lighting variations, occlusions, background movement, and differences in human body structure can affect recognition performance. Real-time processing also requires computational efficiency, which poses constraints on low-resource hardware. Furthermore, generalizing HAR systems for diverse users and environments remains an open research problem that continues to drive innovation in this field.

IV. ANALYSIS AND DISCUSSION

The analysis of existing Human Activity Recognition (HAR) systems reveals that substantial progress has been made in the development of accurate, efficient, and real-time activity monitoring models. However, the effectiveness of a HAR system largely depends on the algorithms, feature extraction techniques, and data preprocessing methods employed. This section discusses the comparative performance, strengths, and limitations of traditional, machine learning, and deep learning-based HAR systems, with emphasis on the role of OpenCV in enhancing performance and adaptability [6].

A. Comparative Analysis of HAR Approaches

Traditional HAR systems primarily relied on handcrafted features such as optical flow, background subtraction, and contour analysis. These methods are computationally lightweight and can be implemented easily using OpenCV; however, they are highly sensitive to lighting variations, occlusions, and camera placement. Machine learning models such as Support Vector Machines (SVM) and Decision Trees improved classification performance by learning from labeled feature sets, but their accuracy still depended heavily on feature selection and preprocessing quality [7].

In contrast, deep learning models, especially Convolutional Neural Networks (CNN) and hybrid CNN-LSTM architectures, have significantly outperformed traditional methods by learning spatial and temporal features automatically from raw video frames. Integrating these models with OpenCV-based preprocessing ensures efficient frame segmentation, skeleton detection, and noise reduction, leading to improved recognition accuracy and real-time capability. Empirical studies indicate that hybrid deep learning models achieve average accuracy rates between 92–96%, whereas classical approaches remain between 80–88% under similar conditions [6].

B. Role of OpenCV in Performance Enhancement

OpenCV serves as the backbone of the majority of HAR implementations due to its optimized libraries for image processing and motion analysis. It enables fast frame acquisition, feature extraction, and object detection, reducing computational overhead for machine learning models. Moreover, OpenCV facilitates integration with external libraries such as TensorFlow and PyTorch, allowing hybrid pipelines that combine traditional preprocessing with deep learning classification. This synergy enhances real-time response and accuracy in dynamic environments such as hospitals or homes where rapid decision-making is critical [5].

C. Computational and Real-Time Considerations

Real-time performance remains a critical requirement for HAR systems, especially in healthcare applications like fall detection. Traditional OpenCV-based systems are faster but may compromise accuracy, while deep learning models deliver better precision at the cost of higher computational power. The adoption of GPU acceleration and lightweight neural networks, such as MobileNet and EfficientNet, offers a viable solution to maintain both speed and accuracy. Additionally, preprocessing using OpenCV—such as background subtraction and ROI extraction—reduces unnecessary data load, ensuring faster model inference and improved system responsiveness.

D. Challenges and Limitations Identified

Despite notable advancements, several challenges persist. Variations in illumination, occlusions, and background complexity still affect detection reliability. Moreover, dataset diversity is often limited, leading to models that perform well in controlled environments but struggle in real-world scenarios. Another key limitation is system scalability—extending HAR frameworks to multiple cameras or remote locations demands efficient synchronization and data management. Addressing these challenges requires improved generalization techniques, domain adaptation, and robust data augmentation strategies.

V. FUTURE RESEARCH DIRECTIONS

The analysis highlights the potential of combining OpenCV's image processing capabilities with adaptive deep learning models for enhanced activity recognition. Future research should focus on hybrid models that fuse visual data with sensor inputs, enabling multimodal HAR systems. Furthermore, edge computing and IoT integration can minimize latency and make real-time HAR feasible even on low-power devices. Explainable AI approaches could also improve transparency in HAR decision-making, increasing trust and usability in healthcare and surveillance systems.

VI. CONCLUSION

Human Activity Recognition (HAR) has emerged as one of the most transformative areas in computer vision and artificial intelligence, offering vast potential in healthcare monitoring, surveillance, and smart automation systems. This survey explored the evolution of HAR techniques, emphasizing the critical role played by OpenCV in simplifying image processing, motion detection, and feature extraction. Through a detailed review of existing literature, it is evident that HAR systems have evolved from traditional handcrafted methods to highly intelligent deep learning architectures capable of real-time and context-aware analysis. The integration of OpenCV with machine learning and deep learning frameworks such as TensorFlow, Keras, and PyTorch has greatly enhanced the accessibility and performance of HAR solutions. OpenCV's efficiency in preprocessing and feature extraction reduces computational overhead and allows seamless real-time operation, which is essential in healthcare applications such as fall detection and patient monitoring. Furthermore, hybrid approaches combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have demonstrated superior accuracy in recognizing complex human actions from video streams.

Despite significant progress, several challenges persist. HAR systems still face difficulties in dealing with environmental variability, occlusion, illumination changes, and limited dataset diversity. Additionally, the computational demands of deep learning models restrict their deployment on low-power edge devices. Addressing these limitations through lightweight model optimization, multimodal data fusion, and adaptive learning will be crucial in advancing this field further.

In conclusion, Human Activity Recognition using OpenCV stands at the intersection of computer vision, artificial intelligence, and human-centered design. As technology continues to evolve, HAR systems will become more intelligent, reliable, and integral to the development of smart, safe, and adaptive environments.

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 [4].

A. Integration with Internet of Things (IoT) and Edge Devices

One of the most promising future directions is the integration of HAR systems with IoT and edge computing platforms. Deploying HAR algorithms directly on edge devices such as smart cameras or microcontrollers can significantly reduce latency and network dependency. This enables real-time activity monitoring in healthcare facilities, homes, and public areas. Combining OpenCV's lightweight image processing capabilities with IoT sensors can further enhance accuracy by providing multi-modal data such as motion, temperature, and vital signs [3].

B. Improved Dataset Diversity and Generalization

Many current HAR models perform well only under controlled conditions due to limited dataset diversity. Future work should focus on developing large, heterogeneous datasets that include variations in human posture, lighting, background, and camera angles. Synthetic data generation using computer graphics or generative adversarial networks (GANs) can help address data scarcity and improve model generalization across different environments and users [2].

C. Multimodal and Context-Aware Recognition Systems

Next-generation HAR systems are expected to be multimodal—combining video data with other sensor inputs such as accelerometers, gyroscopes, or wearable sensors. Context-aware HAR models could adapt to user behavior, surroundings, and activity context, providing intelligent decision-making capabilities. For example, a healthcare HAR system could distinguish between normal lying down and accidental falling by combining motion data with contextual cues such as time, heart rate, or patient activity logs [3].

D. Use of Explainable AI (XAI) in HAR

As HAR systems increasingly influence critical domains like healthcare and security, transparency in decision-making becomes essential. Future research should incorporate Explainable AI (XAI) frameworks to provide visual or textual explanations for model predictions. This will help users and medical professionals understand why a certain activity was detected, thereby improving trust and accountability in automated systems [4].

E. Optimization for Real-Time and Low-Power Devices

Although deep learning-based HAR systems provide high accuracy, they often require significant computational resources. Future work must focus on model optimization through techniques like pruning, quantization, and

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 5, Issue 1, November 2025

Impact Factor: 7.67

knowledge distillation to make HAR systems more efficient on low-power devices. Integration with OpenCV's hardware acceleration and GPU support will further enhance real-time performance without compromising accuracy [6].

Expansion Toward Human Emotion and Gesture Recognition:

Beyond basic activities like walking or sitting, future HAR systems could extend to recognizing human emotions and complex gestures. Combining facial expression analysis with body movement tracking using OpenCV and deep learning can create emotionally intelligent systems capable of interacting naturally with humans. Such systems will find applications in mental health monitoring, robotics, and adaptive user interfaces [5].

REFERENCES

[1]. J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep Learning for Sensor-Based Human Activity Recognition: Overview, Challenges and Opportunities," Pattern Recognition Letters, vol. 119, pp. 3–11, 2019.

[2]. S. Bhattacharya and P. Das, "Real-Time Human Activity Recognition Using OpenCV and Convolutional Neural Networks," Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 9, pp. 3571–3582, 2020.

[3]. Krishna, S. B. Siva; Reddy, K. Rusheendra; Satish, V.; Babu, E. Anka; Vinod, G. V. "Human Activity Recognition With OpenCV and Deep Learning." International Journal of Engineering Research & Technology (IJERT), Vol. 13, Issue 02, February 2024.

[4]. Putra, Ichsan Arsyi; Nurhayati, Oky Dwi; Eridani, Dania. "Human Action Recognition (HAR) Classification Using MediaPipe and Long Short-Term Memory (LSTM)." TEKNIK Journal, Vol. 43, No. 2 (2022).

[5]. Host, Kristina; Ivašić-Kos, Marina. "An overview of Human Action Recognition in sports based on Computer Vision." Heliyon, 2022.

[6]. Janapati, Malathi; Allamsetty, Leela Priya; Potluri, Tarun Teja; Mogili, Kavya Vijay. "Gait-Driven Pose Tracking and Movement Captioning Using OpenCV and MediaPipe Machine Learning Framework." Engineering Proceedings / MDPI, 2024.

[7]. Gupta, Sujata; Ramya, K. R.; Karnati, R. "A Review Work: Human Action Recognition in Video Surveillance Using Deep Learning Techniques." Informatics & Automation, Vol. 23, No. 2, 2024.

[8]. Wang, Xiao; Wu, Zongzhen; Jiang, Bo; Bao, Zhimin; Zhu, Lin; Li, Guoqi; Wang, Yaowei; Tian, Yonghong. "HARDVS: Revisiting Human Activity Recognition with Dynamic Vision Sensors." arXiv / Pre-print, 2022.

[9]. R. Singh and A. Gupta, "Human Activity Recognition Using Multi-Camera OpenCV Framework," International Journal of Computer Applications, vol. 183, no. 35, pp. 1–6, 2021