Human Activity Recognition Using Machine Learning Like Running,Walking,Boxing And Clapping

Shubham U M[1], Varsha A H[2], Prof. Nethravathy V[3], Sowmya M[4], Chinthan M I[5]
[1,2,4,5] Department of CSE Bangalore Institution of Technology Bangalore, India
[3]Assistant Professor Bangalore Institution of Technology Bangalore, India

***Abstract*— Human action recognition is a vital area in computer vision with applications in security, healthcare, and robotics. This paper presents a robust hierarchical method to extract motion information while minimizing the impact of background movements. The proposed approach begins with segmenting the subject from video frames, followed by Motion History Image (MHI) computation to capture motion patterns over time. Significant features are extracted using the Blob Algorithm, which identifies regions of interest. These features are then classified using Support Vector Machine (SVM) and Gaussian Mixture Model (GMM) techniques. The combination of SVM's discriminative power and GMM's probabilistic framework ensures accurate prediction and classification of human actions. Experimental results validate the effectiveness of the approach, showcasing its capability to handle complex scenarios with background noise, making it a promising solution for action recognition in dynamic environments.**

***Keywords*— Human Action Recognition, Computer Vision, Motion History Image (MHI), Blob Algorithm, Support Vector Machine (SVM), Gaussian Mixture Model (GMM), Feature Extraction, Background Noise Reduction, Dynamic Environments, Motion Analysis**

I. INTRODUCTION

Human Activity Recognition (HAR) is a crucial field that focuses on recognizing and categorizing human behaviors from data, including photographs, sensor readings, and video recordings. As a result of the quick development of technology, HAR has become a crucial component of many applications, including as sports analytics, human-computer interaction, healthcare monitoring, and security surveillance. To identify activities ranging from simple motions like walking or running to complex behaviors like group interactions or sports activities, motion data must be captured, pertinent features must be extracted, and classification techniques must be used. Even with great advancements, there are still issues, especially when it comes to managing background noise, illumination changes, occlusions, and the variety of human motions.

This paper addresses these challenges by proposing a hierarchical approach leveraging Motion History Images (MHI), Blob Algorithm, and robust classification techniques such as Support Vector Machine (SVM) and Gaussian Mixture Model (GMM). The combination of feature extraction and advanced classification ensures improved accuracy, making the method suitable for real-world dynamic environments

Identification and analysis of human activity from data sources, such as live video recordings, is the main goal of human activity recognition (HAR). Because of its many uses in industries including healthcare, sports, security, and human-computer interface, it has attracted a lot of interest. With HAR, human movements and behaviors—from simple ones like walking or running to more intricate ones like dancing or participating in sports—are automatically detected and categorized. Through the use of sophisticated algorithms and methodologies, HAR offers insightful information on motion patterns, improving computers' comprehension and responsiveness to human movements. This area has enormous potential to further automation in dynamic contexts, improve user experiences, and ensure safety.

This project's main goal is to create an effective Human Activity Recognition (HAR) system by utilizing accelerometer and gyroscope sensor data from wearable technology and cellphones. In order to identify activities like walking, running, sitting, and climbing stairs, this entails preprocessing raw sensor data to extract meaningful features and applying machine learning techniques like Gaussian Mixture Models (GMM) and Support Vector Machines (SVM) to capture complex spatial and temporal patterns in activity data. The system will be developed to be scalable, lightweight, and appropriate for deployment on mobile devices by evaluating and improving the performance of different machine learning methods. Furthermore, the system will be included into an intuitive real-time activity identification program, offering useful solutions for security, fitness, and healthcare applications.

In order to categorize behaviors like walking, running, and sitting, this project creates a Human Activity Recognition (HAR) system using sensor data from cellphones. Applications for the system include fall detection and rehabilitation in the healthcare industry, customized workout tracking in the fitness industry, and the identification of anomalous activity patterns in the security industry. For accurate activity recognition, the project uses machine learning algorithms like GMM and SVM, leveraging accelerometer and gyroscope data. To ensure resilience in a variety of settings, training and evaluation are conducted using public datasets such as UCI HAR and WISDM. This lightweight, scalable solution may be easily integrated into wearables and mobile devices to provide intelligent behavior analysis, personalized insights, and real-time activity tracking. Improvements in HAR systems have a major impact on smart device applications, fitness tracking, and healthcare.

II. LITERATURE SURVEY

As stated in [1], Ignatov's 2021 IEEE publication movements as Space-Time Shapes presents a generalized technique created for examining 2D shapes and expands its use to volumetric space-time shapes brought about by human movements. The methodology's time-consuming nature, however, may limit its usefulness in real-world situations.[2]. M. Chen and A. Hauptmann's 2020 IEEE publication, Learning to Localise Objects using Structured Output Regression, shows better outcomes than binary training techniques. Instead of concentrating only on classification accuracy on a training set, it uses a training process that is especially tailored for the localisation goal. The requirement for more actions during training, which may raise resource demands, is a disadvantage of this strategy.[3] Action Detection in Complex Scenes with Spatial and Temporal Ambiguities by A. Kläser, M. Marszałek, and C. Schmid, published in 2021 under IEEE, addresses the challenge of detecting human actions in complex real-world scenes. The authors employ Linear Support Vector Machines (SVM) in conjunction with Multiple Instance Learning (MIL) to train an action classifier. This approach allows the model to handle spatial and temporal ambiguities during both training and testing phases, making it suitable for unpredictable real-world scenarios. The key contribution of this work lies in its ability to effectively classify human actions despite spatial overlaps and temporal variations. However, the methodology has notable drawbacks, such as increased programming complexity, which could make implementation and debugging challenging. This complexity may also raise the barrier for integrating this method into practical applications. [4] The 2022 IEEE publication "Multi-Person Bayesian Tracking with Multiple Cameras" by Ibrahim et al. presents a state-of-the-art technique for tracking numerous people in three dimensions with multiple cameras. The suggested approach works especially well in settings like public areas or monitoring installations because it is made to manage situations with partial field-of-view overlaps. This study's primary contribution is its reliable multi-person tracking system, which uses Bayesian inference to get high accuracy in spite of the difficulties caused by overlapping fields of view. However, the method has serious shortcomings, such as the need for accurate camera calibration and a high computational complexity, which raise installation and operating expenses. This work highlights areas for optimization in real-world deployments while laying a solid foundation for the advancement of multi-camera tracking technology.[5]. *View-Independent Action Recognition from Temporal Self-Similarities"* by P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, published in 2021 under IEEE, introduces a self-similarity-based descriptor for video analysis. The primary focus is on achieving view-independent recognition of human actions, which is a significant step forward for handling diverse viewpoints in video data. The proposed method leverages temporal self-similarities to identify recurring patterns of action, offering a robust framework for analyzing dynamic human behaviors. Despite its innovation, the study highlights a limitation in the accuracy of the classifier, which is relatively low. This shortcoming could restrict its applicability in scenarios where high precision is critical, such as surveillance or autonomous systems. The work sets a foundation for future research to refine and enhance classifier performance while maintaining its view-independence capabilities.

III. PROPOSED METHODOLOGY

We characterize human activities using data collected from motion-detecting devices. The dataset covers a variety of features, including accelerometer and gyroscope measurements taken by specialist sensors. These sensors were used to track the motions and vibrations related with various activities.

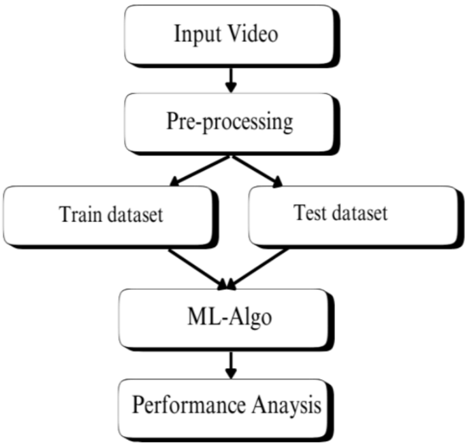

Fig-1 Implemetation Of Model

a. Video Input Source

CCTV footage acquired by strategically placed fixed color cameras in crucial regions serves as the system's input source, allowing continuous monitoring. The video stream captures a variety of human interactions and movements in both public and private places. The video footage's resolution and frame rate are critical components of the recognition pipeline. Higher resolution and frame rate improve the accuracy of object detection and activity recognition by giving more detailed and trustworthy visual input to process.

b. Preprocessing Module

Several preprocessing approaches can be used to improve video analysis and behavior and object detection. By removing random pixel fluctuations like flickering or static noise, noise reduction increases the precision of object tracking and detection. In dimly lit movies, like surveillance footage taken at night, lighting changes balance brightness and contrast levels, increasing the visibility of hidden objects and movements. By stabilizing unstable camera video, motion stabilization makes it easier to focus on the movements of the subject. Furthermore, edge detection methods, such as Canny edge detection, facilitate the recognition process by assisting in the identification of object boundaries.

c. Object Detection and Tracking

Techniques like blob tracking and contour tracking are frequently used to locate and follow people or objects in video footage. Blob tracking tracks and identifies linked areas (blobs) of related pixels that represent moving objects, such as people. . It is perfect for rapidly detecting objects based on motion and is computationally efficient, but it has trouble when items are near or cross one another. Contrariwise, contour tracking is suited for managing intricate shapes and identifying objects in congested environments since it concentrates on identifying and following an object's edges or contours throughout time. Although contour tracking is good at determining the size and shape of objects, it needs to extract features precisely and can have trouble with occlusions, like when a person is momentarily obscured by an object.

d. Behaviour Recognition

To analyze and classify particular behaviors, including walking, applauding, boxing, or hand waving, feature extraction techniques and machine learning algorithms are utilized to identify and categorize these actions. Finding useful information from the video or tracked objects, such as position, acceleration, velocity, and motion trajectories, is known as feature extraction. Understanding acts that take place across time also requires an awareness of temporal aspects, such as the rhythm of hand waving or the periodicity of walking. For categorization, machine learning models such as Support Vector Machines (SVM) and Gaussian Mixture Models (GMM) are frequently utilized. By segmenting the feature space into probabilistic clusters, GMM can model continuous behaviors. This enables the identification of actions such as clapping or walking based on observed data that fits specific distributions. SVM can differentiate between different activities by examining their distinct movement features, such as the range of motion of a hand wave action, after being trained on labeled data. To categorize the behavior, features are extracted and then fed into classifiers such as GMM or SVM. Hand waving is characterized by a back-and-forth motion, boxing by directed hand movements, clapping by quick hand motion, and walking by its periodic movement pattern. By eliminating noise and guaranteeing consistent behavior detection across time, postprocessing techniques like temporal smoothing can improve the classification results.

IV. RESULT

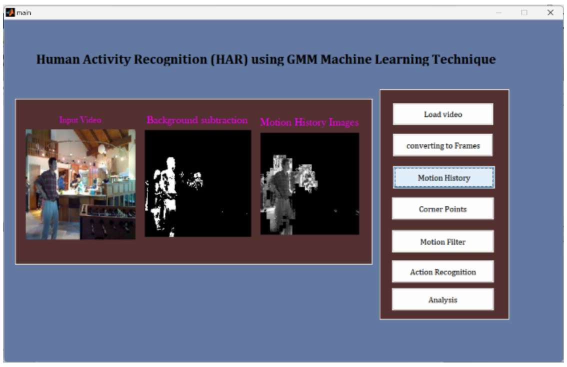

Fig-1. Background subtraction & object detection

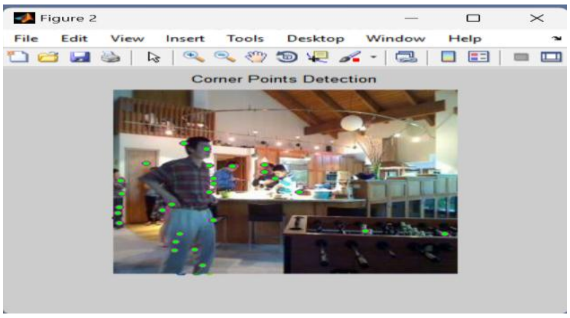

Fig-2. corner points detection

Fig-3. Motion Filter

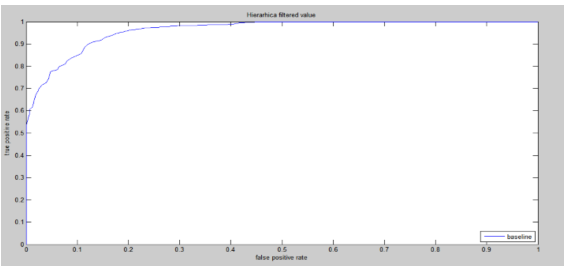

Fig-4. Performance analysis

An essential tool for assessing how well classification models function, especially in applications like activity identification, is the Receiver Operating Characteristic (ROC) curve. It facilitates the analysis of the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR), two important measures. While FPR shows the frequency at which routine actions are mistakenly classified as suspicious, TPR, sometimes referred to as sensitivity, gauges how well the model detects positive cases, such as identifying a particular activity (such as fighting or loitering). The ROC curve, which plots these metrics against various decision thresholds, helps researchers and developers balance the detection of true positives and the reduction of false alarms by revealing how well the model can differentiate between various activities.

An important technique for evaluating the performance of activity identification systems is the ROC curve. For applications that demand accurate identification of anomalous activity, a model with a steep ROC curve around the top-left corner implies high accuracy, with a high TPR and a low FPR. Conversely, a curve nearer the diagonal indicates that the model has trouble differentiating between classes, which leads to a higher number of false positives. Researchers can choose the best threshold for optimal performance by analyzing the ROC curve, which guarantees that the activity identification system satisfies the required accuracy and reliability standards.

V. CONCLUSION & FUTURE SCOPE

This study provides an efficient Human Activity Recognition (HAR) system that combines machine learning models such as Gaussian Mixture Models (GMM) and Support Vector Machines (SVM) with Motion History Images (MHI) and Blob Algorithm for feature extraction. The method improves the accuracy of activity recognition in dynamic situations by addressing issues including occlusions, background noise, and variations in illumination. The system is lightweight and scalable, making it appropriate for wearable and mobile devices. It uses accelerometer and gyroscope data from publicly accessible datasets such as UCI HAR and WISDM. Applications in the fields of healthcare (fall detection and rehabilitation), security (anomaly detection), and fitness (individualized workout tracking) have greatly advanced the automation, user experience, and safety of HAR systems across several industries.

REFERENCES

[1] C. Fernandez, P. Baiget, X. Roca, and J. Gonzalez, "Interpretation of complex situations in a semantic-based surveillance framework," Image Commun., vol. 23, no. 7, pp. 554–569, Aug. 2008.

[2] Dr. H S Mohan and Mahanthesha U, "Human action Recognition using STIP Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-7, May 2020

[3] Y. Changjiang, R. Duraiswami, and L. Davis, "Fast multiple object tracking via a hierarchical particle filter," in Proc. 10th IEEE ICCV, 2005, vol. 1, pp. 212–219.

[4] Loza, W. Fanglin, Y. Jie, and L. Mihaylova, "Video object tracking with differential Structural SIMilarity index," in Proc. IEEE ICASSP, 2011, pp. 1405–1408.

[5] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 5, pp. 564–577, May 2003.

[6] Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. ACM SigKDD Explor. Newsl. 2011, 12, 74–82.

[7] Ignatov, A. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. Appl. Soft Comput. 2018, 62, 915–922 Shwetha K, Spoorthi M, Sindhu S S, Chaithra D. Breast Cancer Detection Using Deep Learning Technique. IJERT,24-04-2018.DOI: 10.17577/IJERTCONV6IS13102.

[8] Sousa, W.; Souto, E.; Rodrigres, J.; Sadarc, P.; Jalali, R.; El-Khatib, K. A comparative analysis of the impact of features on human activity recognition with smartphone sensors. In Proceedings of the 23rd Brazillian Symposium on Multimedia and the Web, Gramado, Brazil, 17–20 October 2017; pp. 397–404. Anvit M, Surabhi S N (2019).

[9] Herna´ndez, F.; Sua´rez, L.F.; Villamizar, J.; Altuve, M. Human Activity Recognition on Smartphones Using a Bidirectional LSTM Network. In Proceedings of the XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), Bucaramanga, Colombia, 24–26 April 2019; pp. 1–5..