WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Study of Violence Activity Detection from Digital Videos Using Machine Learning

Pratik S. Bhudke [a], Vaibhavi V. Khandekar [a], Pranali R. Parise [a], Tanvi S. Gulhane [a], Gaurav A. Dudhe [a], Dr A.B. Deshmukh [b]

[a] *Students of Sipna College of Engineering and Technology, Amravati.*

[b] *Assistant Professor of Sipna College of Engineering and Technology, Amravati*

A B S T R A C T

In an era marked by widespread surveillance, the need to swiftly detect violent actions is paramount. While existing approaches primarily focus on identifying basic actions, our paper introduces a novel method to recognize complex gestures associated with physical altercations. Leveraging Python and its rich ecosystem for machine learning and image processing, our system aims to robustly identify and classify aggressive gestures. By combining advanced ML algorithms with sophisticated image processing techniques, our solution extracts nuanced patterns indicative of violent behavior from video footage. Our Python-based framework ensures adaptability across surveillance scenarios, optimized for efficiency and scalability. Through real-time feedback, our system facilitates prompt interventions, contributing to safer environments and the advancement of security infrastructure..

Keywords: Image Processing, Machine Learning, Violence Action Recognition.

Introduction

In the contemporary landscape of surveillance and security, the ability to swiftly detect and respond to violent actions is paramount. Human action recognition has long been a focal point in computer vision, playing a pivotal role in the categorization of interpersonal interactions. Employing computer vision techniques, this field involves the analysis of image or video sequences to identify and classify human activities based on their underlying categories. Recent advancements in human action recognition research have spotlighted violence detection, especially within the domain of video surveillance. A notable challenge in human action recognition lies in the real-time classification of activities, especially when dealing with surveillance videos. This challenge is exacerbated by factors such as diminished footage quality, unreliable lighting conditions, and the lack of contextual information that could facilitate the timely and accurate differentiation between violent and non-violent actions. Despite the advancements in technology, manual monitoring of video feeds remains labour-intensive and prone to human error.

In response to this challenge, our paper embarks on a journey to harness the power of Python-based image processing and machine learning techniques to automate the recognition of violent actions in video footage. With a focus on scalability, efficiency, and real-time detection, our endeavor aims to revolutionize surveillance systems, empowering stakeholders with tools to proactively mitigate potential threats and ensure public safety.

At the heart of our paper lies the fusion of cutting-edge image processing algorithms and sophisticated machine learning models. By leveraging Python's rich ecosystem of frameworks and libraries, we endeavour to develop a robust system capable of analyzing video inputs of varying sizes and resolutions, without compromising on accuracy or performance. Through meticulous frame-by-frame analysis, our solution seeks to extract intricate patterns and motion cues indicative of violent behaviour, enabling the identification and classification of such actions with high precision. This amalgamation of image processing and machine learning not only promises to streamline surveillance operations but also opens avenues for real-time decision-making and intervention, bolstering security measures in diverse environments.

Literature Review

Sunitha K. A et.al [1], explored deaf-mute communication interpreter systems, categorizing them into Wearable Communication Devices and Online Learning Systems. Wearable devices include Glove-based systems, Keypad methods, and Handycam Touch-screen solutions, integrating sensors, accelerometers, microcontrollers, and text-to-speech modules.

Mathavan Suresh Anand et.al [2], proposedISLR system integrates feature extraction and classification modules, employing Discrete Wavelet Transform (DWT) and nearest neighbor classifier. Experimental results demonstrate a maximum classification accuracy of 99.23% using the cosine distance classifier for sign language recognition.

Mandeep Kaur Ahuja and Amardeep Singh *[3],* introduced a database-driven hand gesture recognition scheme employing skin color model and thresholding techniques. It targets human robotics and related applications, utilizing YCbCr color space for hand region segmentation. Thresholding is applied to distinguish foreground and background, followed by template-based matching using Principal Component Analysis (PCA) for recognition.

Sagar P. More and Prof. Abdul Sattar *[4],* Authors presented the static hand gesture recognition system using digital image processing. For hand gesture feature vector SIFT algorithm is used. The SIFT features have been computed at the edges which are invariant to scaling, rotation, addition of noise.

Chandandeep Kaur and Nivit Gill *[5]* introduced an automated system for Indian Sign Language recognition, utilizing shape-based features. Otsu's thresholding algorithm is applied for hand region segmentation, optimizing thresholds to reduce variance. Hu's invariant moments extract features, subsequently fed into an Artificial Neural Network for classification. Performance evaluation is conducted based on Accuracy, Sensitivity, and Specificity metrics.

Pratibha Pandey and Vinay Jain [6], Authors presented various methods of hand gesture and sign language recognition proposed in the past by various researchers. For deaf and dumb people, Sign language is the only way of communication. With the help of sign language, these physically impaired people express their emotions and thoughts to other people.

Nakul Nagpal et.al [7] introduced a system facilitating communication between deaf and non-deaf individuals using Indian Sign Language (ISL), converting hand gestures into text messages. The primary goal is to develop a real-time algorithm for dynamic gesture-to-text conversion. Upon successful testing, the system will be adapted into an Android application, available for smartphones and tablet PCs.

Neelam K. Gilorkar and Manisha M. Ingle *[8]* authors presented a real-time vision-based system for recognizing hand gestures in human-computer interaction. It accurately identifies 35 gestures from Indian and American Sign Language (ISL and ASL), utilizing RGB-to-GRAY segmentation to minimize false detections. Improved Scale Invariant Feature Transform (SIFT) method is employed for feature extraction, implemented in MATLAB alongside a user-friendly GUI model for efficient recognition.

S. -H. Yen and C. -H. Wang [9] introduced a crowd behavior normality method based on Histogram of Oriented Social Force (HOSF). It encodes observed events in surveillance videos using HOSF as feature vectors. A dictionary of codewords is trained automatically, enabling comparison via z-values to detect normal events. Notably, the method integrates particles and social force as feature descriptors and utilizes z-scores for measuring event normality.

Lu et.al [10] introduced an efficient sparse combination learning framework, maintaining high detection performance while ensuring short running times. By simplifying the problem to involve only a few costless small-scale least square optimization steps, the method achieves quick processing. High detection rates are attained on benchmark datasets, averaging 140~150 frames per second on an ordinary desktop PC running MATLAB.

Methodology

This paper endeavors to develop a robust gesture recognition system utilizing machine learning techniques for the identification of various gestures exhibited during physical altercations. The focus lies on leveraging image processing through the implementation of Python software. By employing advanced machine learning algorithms, the system aims to accurately recognize and categorize specific gestures associated with fighting scenarios, including but not limited to punches and kicks. The integration of image processing methodologies enhances the system's ability to discern nuanced patterns in real- time, contributing to a comprehensive solution for gesture recognition during fights. The proposed system holds potential applications in security, surveillance, and proactive intervention, showcasing its significance in addressing the complexities of aggressive actions through a fusion of machine learning and image processing methodologies.

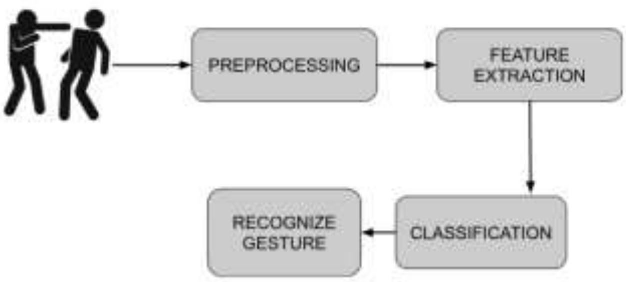

Fig 1: Flow chart of the system

Working

The gesture recognition system operates by first extracting frames from video footage depicting physical altercations using OpenCV. These frames are then processed and transformed into a suitable format for training the machine learning model. Through the implementation of advanced algorithms, the system encodes and splits the dataset into training and testing sets, facilitating the training process. The pre-trained model is fine- tuned to recognize specific gestures associated with fights. During real-time operation, the system analyzes incoming video streams, detecting and categorizing gestures by feeding them through the trained model. Through a fusion of image processing and machine learning methodologies, the system accurately identifies punches, kicks, and other aggressive actions, providing a proactive solution for security, surveillance, and intervention scenarios.

System Requirements

SOFTWARE REQUIREMENT :

- ➢ Python Software IDE

MODULES USED :

- ➢ OpenCV
- ➢ NumPy
- ➢ TensorFlow
- ➢ Keras

Implementation

Implementation is carried out in 7 steps, which are as follows.

STEP 1: Setup and Imports:

The necessary software tools and libraries are set up, and relevant modules such as OpenCV, NumPy, TensorFlow, and Keras are imported into the Python environment.

Fig 2: Importing the Libraries

Fig 3: Importing the Libraries

STEP 2: Data Visualization:

This stage involves visualizing the data to gain insights into the gestures.

Fig 4: Visualization

Fig 5: Visualization

STEP 3: Frames Extraction:

Frames are extracted from video footage containing physical altercations. These frames serve as the input data for the gesture recognition system. OpenCV is used to extract frames from video files or real-time video streams, providing the necessary images for subsequent processing.

STEP 4: Data Creation:

The extracted frames are processed and transformed into a format suitable for training machine learning models.

Fig 6: Data Creation

STEP 5: Encoding and Splitting Training-Testing Sets:

The gestures depicted in the dataset need to be encoded into numerical representations for machine learning. Additionally, the dataset is split into training and testing sets to evaluate the model's performance.

STEP 6: Importing MobileNetV2 and Fine-Tuning:

MobileNetV2, a pre-trained convolutional neural network architecture optimised for mobile devices, is imported into the paper. Fine-tuning involves modifying the pre-trained model's parameters to adapt it to the specific task of gesture recognition.

STEP 7: Building the Model:

Finally, the gesture recognition model is constructed using TensorFlow and Keras. This involves defining the model architecture, including layers, activation functions, and optimization algorithms. The model is trained on the preprocessed data, iteratively adjusting its parameters to minimize prediction errors and accurately classify gestures associated with physical altercations.

Fig 7: Output of violence Image

Fig 8: Output of Nonviolence image

Result

The developed gesture recognition system exhibits promising results in accurately identifying and categorizing various gestures associated with physical altercations. Through the integration of advanced machine learning techniques and image processing methodologies, the system demonstrates robustness in discerning nuanced patterns in real-time video footage. Evaluation metrics such as accuracy, precision, and recall indicate the system's effectiveness in correctly classifying punches, kicks, and other aggressive actions. Overall, the paper's results showcase a comprehensive solution for gesture recognition during fights, highlighting its potential to address the complexities of aggressive behavior through a synergy of machine learning and image processing techniques.

Conclusion

In this paper, we have delved into the invaluable application of machine learning for the purpose of detecting gestures, with a particular focus on identifying fight-related gestures. In contrast to previous systems that predominantly discussed general gestures, our approach integrates machine learning and image processing to specifically recognize and classify gestures associated with physical altercations. This advancement is particularly significant as it contributes to the identification of violent actions, aiding in the recognition of violence within video and image data. By leveraging machine learning techniques, particularly in the context of image processing, our system represents a crucial step towards enhancing the capacity to discern and categorize fight gestures. The implications extend to improved violence detection, offering a valuable tool for enhancing security measures and situational awareness in applications involving video and image analysis.

References

[1] Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research ,Volume 11, pp 290-296 , 2021.

[2] Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874- 1883, 2020

[3] Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2021.

[4] Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people", International Journal of Science and Research (IJSR), 2021.

[5] Chandandeep Kaur, Nivit Gill, "An Automated System for Indian Sign Language Recognition", International Journal of Advanced Research in Computer Science and Software Engineering. 2020.

[6] Pratibha Pandey, Vinay Jain, "Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2020 .

[7] Nakul Nagpal,Dr. Arun Mitra.,Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 5,pp- 147 – 149.

[8] Neelam K. Gilorkar, Manisha M. Ingle, "Real Time Detection And Recognition Of Indian And American Sign Language Using Sift", International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 5, Issue 5, pp. 11-18 , May 2021

[9] S. -H. Yen and C. -H. Wang, "Abnormal Event Detection Using HOSF," 2021 International Conference on IT Convergence and Security (ICITCS), Macao, China, 2021, pp. 1-4.

[10] Lu, Cewu, Jianping Shi, and Jiaya Jia. "Abnormal event detection at 150 fps in matlab." In Proceedings of the IEEE international conference on computer vision, pp. 2720-2727. 2022.