IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Suspicious Human Activity Detection

Mr. Chauhan Rahul B [1], Mr. Nimesh Vaidya [2] , DR.Vijay Gadhavi [3]

[1]PG Scholar – Faculty of Engineering, Computer Engineering Department Swaminarayan University, India

[2]Assistant Professor & HOD - Faculty of Engineering, Computer Engineering Department Swaminarayan University, India

[3]Associate Professor & Dean –Faculty of Engineering (I/C), Computer Engineering Department Swaminarayan University, India

Abstract

Human activity detection for video surveillance system is an automated way of processing video sequences and making an intelligent decision about the actions in the video. It is one of the growing areas of Computer vision and artificial intelligence. Suspicious Human Activity Detection is the process of detecting unwanted human activities in a crowded place and alerting the concerned authority about the activity. This is done by converting video into frames and analyzing the persons and their activities from the processed frames. This paper gives an overall idea of the development of the system that detects suspicious human activities.

Keywords

Suspicious crowd behavior, Video surveillance, Transfer learning, Anomaly detection, Computer vision

Introduction

The widespread incorporation of many applications in modern society has significantly transformed many aspects of our lives, with visual systems emerging as essential instruments. One important area of study in this field is the detection of suspicious human behaviour using video surveillance, which involves classifying behaviours as either normal or abnormal [1]. The increasing frequency of disruptive incidents in public areas globally, ranging from banks to airports, highlights the urgent requirement for efficient security measures [2]. As a result, surveillance systems, mostly dependent on CCTV cameras, have grown quite common, producing large quantities of video data for examination. Nevertheless, the labour-intensive nature of manual monitoring makes it unfeasible, thus necessitating the development of automated detection systems [3].

A. Research Contribution

- Presented a novel approach for detecting potentially suspicious human behaviour by leveraging deep learning techniques. The efficacy of the suggested methodology is validated by conducting prediction scenarios using previously unseen test data and by creating a YouTube video.
- By utilizing Convolutional Neural Networks (CNNs) and sophisticated deep learning structures, such as the suggested Time Distributed CNN model and Conv3D model, we attain notably enhanced accuracy rates of 90.14% and 88.23%, respectively, surpassing current research approaches.

The following sections of this paper are organized to provide a more detailed examination of the background and relevant literature (Section II),

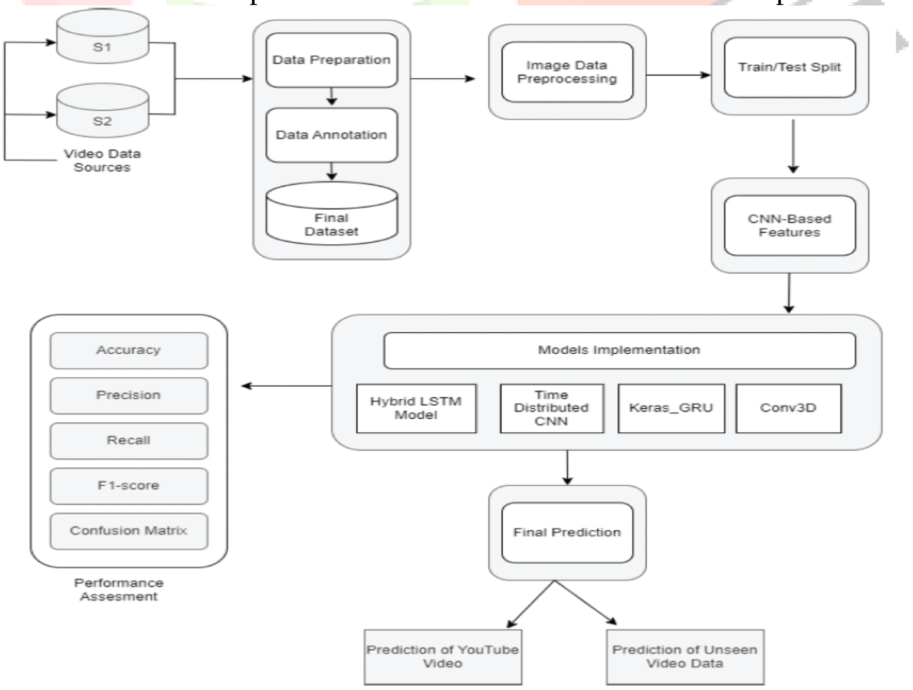
Literature Review

While Human activity recognition has been a topic of considerable study in present literature, this section delves into the latest improvements in this domain. Cutting-edge studies in Human activity recognition predominantly revolve across the realms of machine learning and deep learning methodologies.

In the area of machine learning, Ghazal et al. [10] carried out a comparative study specializing in Human activity recognition with the use of 2d-skeletal facts. They utilized the OpenPose package to extract visual and motion attributes from 2d landmarks of human skeletal joints. The examine evaluated five supervised machine learning strategies, consisting of support Vector machine (SVM), Naive Bayes (NB), Linear Discriminant (LD), k-nearest neighbours (KNNs), and feed-forward backpropagation neural networks. The primary objective was to identify four awesome activity instructions: sitting, standing, walking, and falling, with the k-nearest neighbours (KNNs) exhibiting the most promising overall performance. In another study

Proposed Approach

The proposed methodology for identifying suspicious human activity entails several crucial stages, as shown in Figure 1. Initially, data information is gathered from distinct sources, denoted as S1 and S2. This data then undergoes meticulous preparation regarding cleansing, formatting, and integration to produce a cohesive dataset that consolidates relevant information from both sources. Significant interest is given to the preparation of images within the dataset. Techniques consisting of normalization, scaling, and augmentation are employed to ensure uniformity and enhance the excellent image inputs for the next evaluation. Furthermore, the dataset is annotated, with instances of suspicious human behaviour being categorized to facilitate accurate classification and prediction by using supervised learning algorithms. To facilitate the training and assessment of the model, the dataset is partitioned into separate units for training and testing functions. This ensures that the model is skilled in an extensive extent of data while retaining a distinct subset of data for rigorous testing and evaluation. This division guarantees the integrity of the evaluation process by evaluating the performance of the model on data that has not been previously viewed. CNNs are then utilized to extract significant characteristics from the preprocessed image data. These characteristics act as informative representations of the input data and are crucial in later model implementations.

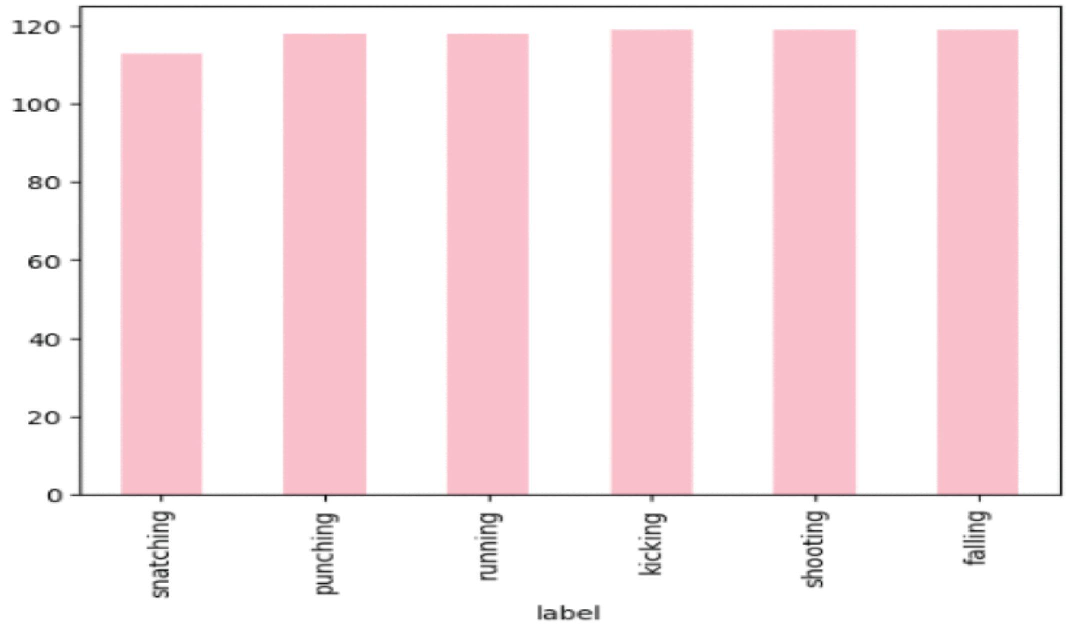


Several deep learning architectures, such as the Hybrid LSTM Model, Time Distributed CNN, Keras_GRU, and Conv3D, are used to tackle the task of identifying suspicious human activity. Every model utilizes the extracted CNN-based features to acquire knowledge of patterns and provide predictions. Ultimately, the trained models are implemented to predict potentially dubious human behaviours in real-life situations, encompassing the anticipation of YouTube videos and the examination of unfamiliar video data. This step highlights the practical usefulness and ability to be applied in various situations involving the established strategy. In general, this methodology seeks to improve surveillance and security systems by increasing the ability to identify and stop possible threats.

Dataset Selection

The dataset component of the research paper outlines the methodology for developing a dataset customized specifically for video classification.

After obtaining the videos, they were sorted into separate files based on their assigned categories. As an illustration, movies that showed falling behaviour were organized in a folder labelled "falling," whereas recordings that displayed kicking activities were placed in a folder titled "kicking." This organization enabled the methodical administration and retrieval of videos during the following phases of data preparation and model training.



A. Dataset Annotation

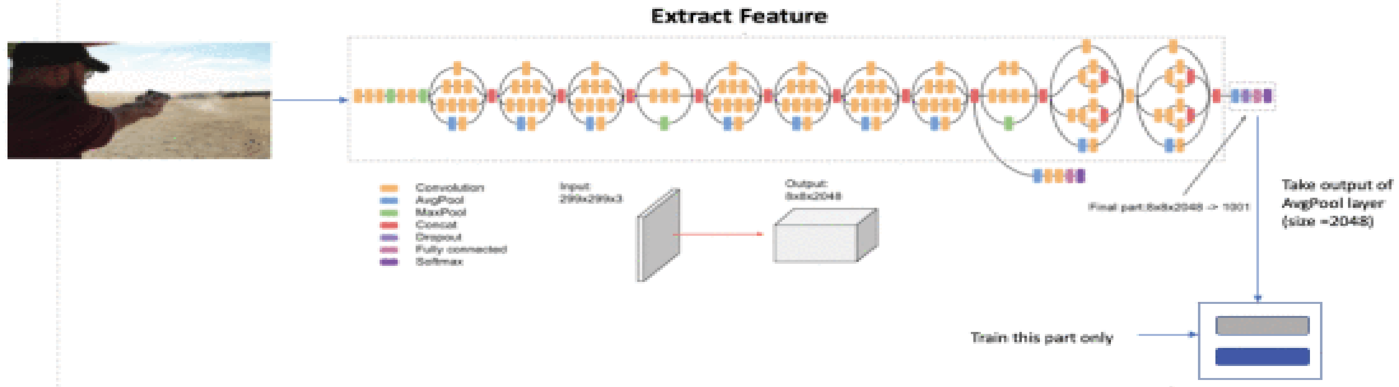
The process of data annotation was automated using Python code to enhance efficiency. The code sequentially scans all video files stored in the specified directory. Afterwards, it takes and stores the name of each video in a cache while also adding this information to a CSV file.

B. Data Pre-Processing

Data preprocessing serves as a foundational step in research implementation, especially in tasks related to video record evaluation. In this study, video information was preprocessed using the Python OpenCV library.

C. Feature Extraction

For feature extraction, we harnessed the power of the InceptionV3 model, a custom variation of CNNs famous for its excellent performance in image analysis responsibilities. Our adaptation of the InceptionV3 model integrated specific parameters tailored to our requirements.



D. Models Implementation

Detecting suspicious human activities inside video recordings relies heavily on the robustness of deep learning architectures. Leveraging improvements in neural network structures, we explore a spectrum of models meticulously designed to capture the dynamic spatio-temporal patterns inherent in video sequences. This section entails a comprehensive examination of the implementation intricacies of high-quality deep learning architectures, together with the Hybrid LSTM

Results

A. Experimental Configuration

The analysis utilized Python 3.8 and the Kaggle IDE. Given the time-consuming process of training deep learning models, it is crucial to ensure that the library installation is done correctly in order to achieve successful model training and execution. TensorFlow is highly regarded as one of the most commonly utilized libraries for creating effective image-processing models. The necessary libraries installed for this experiment comprised TensorFlow, Keras, Scikit-learn, Matplotlib, Pillow, and OpenCV. Keras functions as a leading tool for creating deep learning models and functions as a library with a high-level interface for the TensorFlow framework. By utilizing Scikit-learn, a Python toolbox, one may effectively employ machine learning techniques such as classification and regression. The Matplotlib library in Python is quite useful for visualizing data. Both OpenCV and Pillow are utilized for image-processing jobs. It is crucial to install all of these libraries in order to experiment successfully successfully.

B. Experimental Results

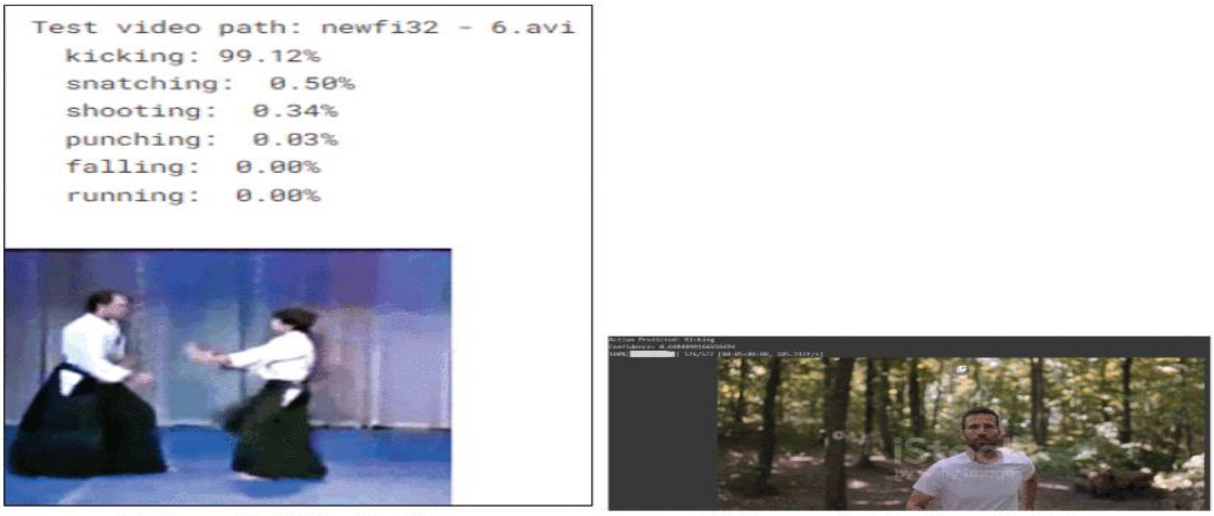
Displays the numerical outcomes derived from assessing several models on the dataset. The performance measures for each model, such as accuracy, precision, recall, and F1-score, are provided.

TABLE 2 Comparative Results for SHAR

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Hybrid Model	84.51	84.89	83.10	84.72
Time Distributed CNN	90.14	90.78	90.14	90.14
Keras_GRU	83.80	87.10	84.10	84.10
Conv3D	88.23	88.20	88.20	88.20

C. Model Prediction

Model prediction entails utilizing a well-trained deep-learning model to produce accurate predictions or classifications on novel, unobserved data. Within this particular framework, the initial stage entails making predictions with the trained model using test data that has yet to be previously encountered.



```
Test video path: newfi32 - 6.avi
   kicking: 99.12%
   snatching:  0.50%
   shooting:  0.34%
   punching:  0.03%
   falling:  0.00%
   running:  0.00%
```

```
Action Predicted: Kicking
Confidence: [illegible]
100%|[illegible]
```

(a) Unseen Test Video Prediction

(b) Unseen Youtube Video Prediction

D. Comparative Analysis

The comparison table, as depicted in Table 3, gives a complete assessment of both the proposed and existing research methodologies within a particular area. The cutting-edge methodology, as mentioned via Khan et al. citekhan2022human, encompasses various models, which include MLP (Multi-Layer Perceptron), CNN, LSTM, BiLSTM (Bidirectional LSTM), and CNN-LSTM, with every model's accuracy provided as a percentage.

	Model	Accuracy (%)
Existing Approach [25]	MLP	71.51
	CNN	75.47
	LSTM	66.09
	BiLSTM	66.26
	CNN-LSTM	76.50
Proposed Approach	Hybrid Model	84.51
	Time Distributed CNN	90.14
	Keras_GRU	83.80
	Conv3D	88.23

Conclusion and Future Work

This study outlines a scientific method for detecting suspicious human activity through a series of crucial procedural steps. Our research started with the gathering of data from different resources, denoted as S1 and S2. We then meticulously refined this data by getting rid of inconsistencies, standardizing its format, and consolidating it right into a unified dataset. Rigorous image preprocessing ensued regarding normalization, scaling, and augmentation strategies to ensure image uniformity and enhance quality.

Moreover, dataset annotation was undertaken to strengthen the precision of classifying and predicting suspicious human behaviour using supervised learning algorithms. To facilitate model training and evaluation, we partitioned the dataset into separate training and testing sets, ensuring unbiased assessment. CNNs have been employed to extract essential features from the preprocessed image data, which is important for future model implementation. To detect suspicious human activity, diverse deep-learning architectures, including the Hybrid LSTM model, time-dispersed CNN, Keras_GRU, and Conv3D, were explored. Each model leveraged the extracted CNN capabilities to gain insights into patterns and offer predictions. Our findings discovered that the proposed time-disbursed CNN model exeTimecuted an appreciably better accuracy rate of 90.14%, showcasing its efficacy in appropriately detecting suspicious human sports. Similarly, the Conv3D model in our proposed method exhibited full-size development as compared to current techniques, yielding an accuracy of 88.23%. In the end, we implemented the trained algorithms to forecast potentially suspicious human movements in real-world scenarios, which includes predicting the content of YouTube videos and scrutinizing surprising video footage. This sensible validation underscores the utility and relevance of our proposed methodology in bolstering surveillance and security systems by enhancing functionality to identify and mitigate potential threats.

Inside the realm of future research and advancement, several avenues present themselves for exploration and refinement. First of all, expanding the proposed technique to encompass a broader array of datasets, consisting of an extra range of suspicious activities and environmental factors, holds promise for boosting the model's efficacy and adaptableness. Additionally, delving into advanced deep learning architectures and methodologies, along with attention mechanisms and transformer models, can significantly enhance the accuracy and efficiency of activity recognition. Furthermore, integrating real-time facts streaming and processing competencies into the monitoring system ought to permit prompt identity and response to any suspicious activities, thereby improving typical security measures. Participating with domain experts and stakeholders is crucial to toughen the proposed technique's robustness and reliability. This collaboration can aid in refining the annotation process of the research and validating the model's predictions in real-time scenarios. Furthermore, considering the ethical considerations and privacy ramifications related to surveillance systems, forthcoming research needs to prioritize the development of obvious and responsible frameworks for statistics series, storage, and utilization. This necessitates the implementation of sturdy records protection measures and adherence to pertinent regulatory requirements to protect individuals' privacy rights while retaining robust security protocols. By way of addressing these important regions of challenge, future research endeavours can contribute to the evolution of surveillance systems that are both ethically sound and operationally effective.

REFERENCE

[1] Video Analytics, White Paper, AGENT Video Intelligence, July 2016.

[2] A Review on Object Detection in Video Processing, International Journal of u- and e- Service, Science and Technology, Vol. 5, No. 4, December, 2012, Kauleshwar Prasad, Richa Sharmaand, Deepika Wadhwani, BIT, Durg, India.

[3] Suspicious Human Activity Detection from Surveillance Videos, (IJIDCS) International Journal on Internet and Distributed Computing Systems, Vol: 2 No: 2, 2012, Gowsikhaa D, Manjunath, Abirami S, Department of information science and technology, Anna University, Chennai, Tamilnadu.

[4] M. Fahad Khan, Hafiz Adnan Habib, "Video Analytics for Quantitative Employee Performance Evaluation", Canadian Journal on Image Processing & Computer Vision Vol. 1, No. 1, pp. 9-15, February 2010.

[5] Benjamin Maurin, Osama Masoud and Nikos Apanikolopoulos, "Camera Surveillance of Crowded Traffic Scenes", IEEE Computer Society Press Vol. 22,No. 4, pp.16-44, 2010.

[6] Gwang Goo K Lee, Hwan Ka, Byeoung Su Kim, Whoi Yul Kim, Ja Young Yoon and Jae Jun Kim, "Analysis of crowded scenes in Surveillance Videos", Canadian Journal on Image Processing & Computer Vision Vol. 1, No. 1, pp.52-75, 2010.

[7] Paul Viola and Michael J. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision Vol. 57, No. 2, pp. 137–154, 2004.

[8] Suspicious Human Activity Recognition for Video Surveillance System, 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Ahmad Salihu Ben-Musa, Sanjay Kumar Singh, Prateek Agrawal, Department of Computer Science and Engineering, Lovely Professional University, Punjab, India.

[9] Face Detection using SURF Cascade, Jianguo Li, Tao Wang, Yimin Zhang, Intel Labs China.