iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** V **Month of publication:** May 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.53312

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue V May 2023- Available at www.ijraset.com*

# Human Activity Tracking System Using Pose Estimation Technique

Sakshi Atul Gaikwad[1], Shravani Shrikant Sarde[2], Pooja Baloo Khade[3], Prof. Ankita Gupta[4]
[1, 2, 3, 4]*Department of Computer EngineeringAISSMS College of Engineering Affiliated to Savitribai PhulePune University Pune City, India*

***Abstract:*** ***When some incident / any dangerous activity happens in the area the CCTV takes photos & video pieces of evidence of what happened there, but there are only few such system that can capture evidence in real-time and process it to conduct further actions of an incident. It waits for the police to come and resolve the issue and what happens is when police come the evidence are vanished. The drawbacks are that when such an incident happens out of 10 people only 1 or 2 will try to call the police or ambulance but this also takes a lot of time, no quick decisions are taken to resolve the issue until the police gets informed about it. Also, if the accident / suspicious activity happens on road then that road routes get blocked in some cases, the system's model should be accurate enough to observe the situation and by predictions should proceed with the decision-making to help police to resolve the issue smoothly and as early as possible. Our solution for such a problem is to develop a system that uses CCTV surveillance systems with an advanced technological model with which it can capture required evidence. Also, we had developed a system that can recognize gestures of a person to detect if he/she is beating someone in the Realtime footage with the help of CNN, Keras, YOLO, MediaPipe and TensorFlow libraries. Quick notifications are sent to alert the police control rooms. At Control room police will allot the nearby required FRVs clusters by sending notifications to them of incident happened in a particular area.***

***Keywords: Computer Vision, Image Processing, Keras, YOLO, TensorFlow, Convolutional Neural Network, Data Analysis, Human Gestures recognition.***

## I. INTRODUCTION

Traditionally, we depend on the manual interference to detect and alert about suspicious activity detection. Suspicious activity detection and alert system is gainingpopularity due to increase in crime rates. Due to this itbecomes necessary to detect suspicious activities automatically. This system helps in video forensic and surveillance system. Today's surveillance systems involve manual interference to resolve the problem. There is the requirement of human computer handballed network of CCTV which can capture data in real-time and if this data recognizes or detects any suspicious activity at its vision then it must pass an alert to required organization. Reducingmanual mistakes and adapting technologies that can give accurate results without any mistake. There must be a system that must work on 80 percent predictions by machine and 20 percent confirmation on incident to take actions quick. Machine can do repetitive tasks. To provide an efficient improvised solution on a social issue we have implemented the solution based on Convolutional neural network-based analysis. To reduce crime rates and collect more evidence regarding incident / suspicious acts.

## II. RELATED WORK

Dr. Gayashan Kariyawasam's with research on "Suspicious activity detection in surveillance footage" gave us an idea that for testing our project the dataset is required which we can get through movies videos, and we can experiment lively at testing phase. The limitations we noticed in the paper is that they used footages for the data testing. It is good while we're in the testing phase, but it doesn't work in real-time.[1]

Then we understood the process to collect such huge data soour next point of view towards the research was to study the ways to detect objects from the video and process it using models of machine learning and artificial intelligence. There was research conducted by Dr. Sorina Smeureanu posted "Real-Time deep learning method for abandoned luggage detection in video" the Finite State Automata, CNN & StaticObject Detection can be useful to design a model to detect abandoned objects [2]. To perform such experiment, they have gone through the false positive results of the system that used Simple Static object detection where they observedthat the algorithms was having too much false positive values that is the model considered humans and other stationary objects those where in the model for the long timeeven though they were not in the footage from the beginning. So, they used the concept of double static object detection algorithmic strategy for the more accuracy.

©IJRASET: All Rights are Reserved 

The thefts like one person attacking on the other person or there might be quarrel between persons in the middle of the road. To detect such events Dr. Rajib Debnath conducted an experiment on "Automatic visual gun detection carried by a moving person" in which he and his teammates designed the system for moving object detection algorithm where activities must support rotation, scaling and math for detecting the objects. [3].

They had considered their own dataset by using firearm database and computing and intelligent information systems to track and observe his gestures to identify whether he is committing any crime or not.

Dr. Kanchana. V with research on "Multiple car detection, Recognition and tracking in traffic" they used Categories of features in vision is made through SVM.

[4] where we got an idea to make our algorithms work in the same frame with fast accuracy by observing multiple objects at a time not the single object at on time. We studied the way stated by Dr. Rasha Saffarini in their research paper named "Smart system to avoid car accidents" gave us an idea that Sensors can be used for detecting the activities. [5]. Computer Visionalgorithms so that we with the help of OpenCV library from python can be more useful and efficient in the process to implement in while detecting the person, car or any object into the footage.

Dr. Nagarjuna R Vatti in the paper named "Smart Road accident detection and communication system" gave us a way of idea of work that, Alerts can be sent for notifying thepolice and fast response vehicles. [6]

There are certain implementations done using object detectors by Dr. Cheng-jia Wang and team by their research on "Real-Time car detection and driving safety alarm system with google tensor-flow object detection API" gave us an idea to detect any object in video or real-time cameras the GTOD API in python language is helpful.[7] And with the usage of confusion matrix for object detector predicts the incident. Dr. Jing-Hao Sun and team had done research on the segmentation of hand gestures by establishing theskin color model and AdaBoost classifier according to the particularity of skin color for hand gestures, as well as the denaturation of hand gestures with one frame of video beingcut for analysis. In this regard, the human hand is segmentedfrom the complicated background, the real-time handgesture tracking is also realized by Cam-Shift algorithm. Then, the area of hand gestures which has been detected in real time is recognized by convolutional neural network to realize the recognition of 10 common digits. Experiments show 98.3% accuracy.[21]

Dr. Kapitanov Alexander with his team they worked on the dataset contains 552,992 samples divided into 18 classes of gestures. The annotations consist of bounding boxes of hands with gesture labels and mark-up of leading hands.[22]

Dr. Fan Zhang presented a real-time on-device hand tracking solution that predicts a hand skeleton of a human from a single RGB camera for AR/VR applications. This pipeline consists of two models: 1) a palm detector, that is providing a bounding box of a hand to, 2) a hand landmark model, that is predicting the hand skeleton. It is implemented via Media Pipe, a framework for building cross-platform ML solutions. The proposed model and pipeline architecture demonstrate real-time inference speed on mobile GPUs with high prediction quality.[23]

## III. SYSTEM ARCHITECHTURE

Basically, we are going to follow agile methodology for the entire workflow of the project. We have designed this project to resolve the issue that happens often in many areas. So, in the architecture diagram above we have included the exact way of how each module or block of the program is going to interacts and share, store, and processes data on by itself with certain human verification and which everything is based on the internet connectivity for all the users. In future work we have also planned for offline system model based on hardware-software interaction that can work smoothly without having the internet related dependability for the system to process.

The system will initially connect with the active camerasin the area that it is located at. These CCTVs are handled andmonitored by the police control room which uses a Server forstoring and processing of the data. With the help of deep learning and machine learning models we have applied algorithms and models that stores the relevant details of the data / evidence that we gathered from the CCTV. Continuously system will simultaneously blink on the objectthat is kept abandoned from certain time. And overall, the processing and extraction of suspicious events such as fighting, thefts are being detected by our models with accuracy of 75% that we have trained our system with.Camera is acting as an input device which captures visuals ofthe incident. In this way we can have evidence of the incident that has happened.

The following fig.3.1 has a whole architecture that we have designed and implemented accordingly. Each module has its own role, functionality, and processing.

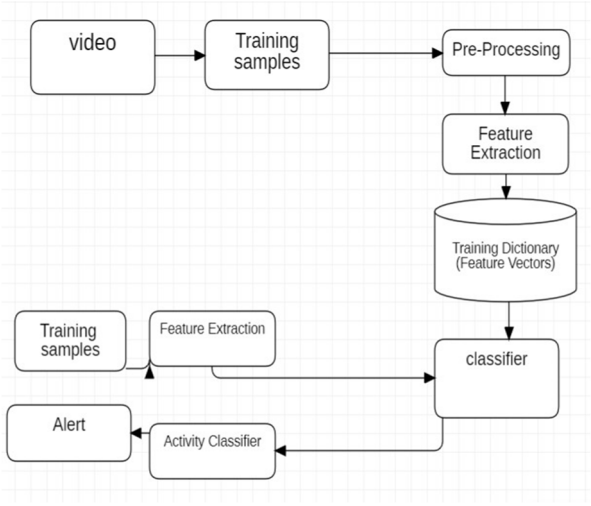



Fig 3.1:Overall system architecture.

## IV. MODULES

We have designed our system in 4 modules, and each module is having it's own specific task like a task of alert module is to send the notification to the police from the system and so on. Given below are the four modules which we had implemented in our project:

1) *Video Capturing Module:* This module is the initial step where we are using camera as our input device. Each frame iscaptured by the camera based on the FPS of camera. Then it is sent to the next module for the further process.
2) *Pre-processing and Recognition Module:* Pre-processing is our second step where our actual model starts its working by recognizing the object and human actions. We have used many training samples to train our model when the model detects the activity it is checked from the trained data to accurately predict the actions of human forfurther classification of the suspicious activity. The output result of pre-processing and recognition module is sent as an input to the next module which is the storage module.
3) *Storage module:* The Storage module takes the input from pre- processing and recognition module and differentiates the action from frame into suspicious or normal. Based on which it stores theframe image in its recognized category.
4) *Alert Module:* Meanwhile, the alert module is initialized as soon as the pre-processing and recognition module detects the suspicious activity performed by a person. And alert is generated as soon as possible and sent to the police control room.

## V. IMPLEMENTATION

By using Google Teachable we had implemented our solutions. In this model we have created separate classes for different activities. We had First create an ML model having datasets of normal activities such as walking, talking, reading, sitting etc. Then feed the datasets of Suspicious Activity such as fighting, boxing, pointing guns or any other violent movement deemed suspicious into the ML model. Also2: Feeding perform such kind of activities in front of the smart camera so that various movements are captured. This will be useful for training the ML model and deploying it to make a smart AI camera.

To create and train an ML model, there are severalflexible options such as TensorFlow, Google Teachable,Edge Impulse, Lobe etc. We the use of Google Teachable, we selected the Pose Net option for tracking the various body movements and actions. Earlier, we performed different actions like walking, talking, eating, standing etc. By correctly labelling them, then we feed these datasets into the ML model. Similarly, wealso feed the datasets of activities like pointing guns, firing guns, fighting, beating etc.

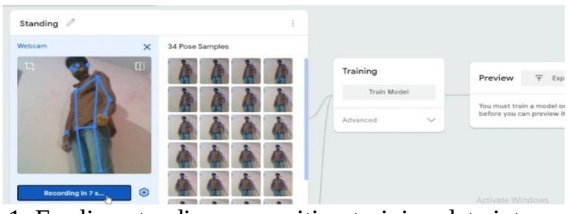



Fig 1: Feeding standing recognition training data into model

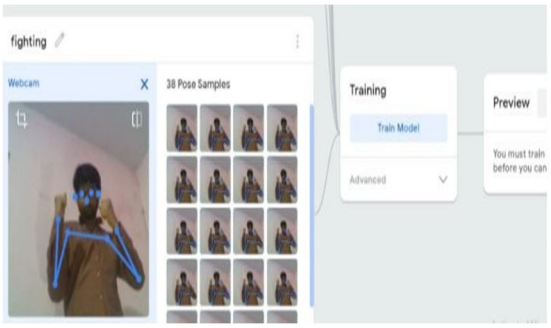


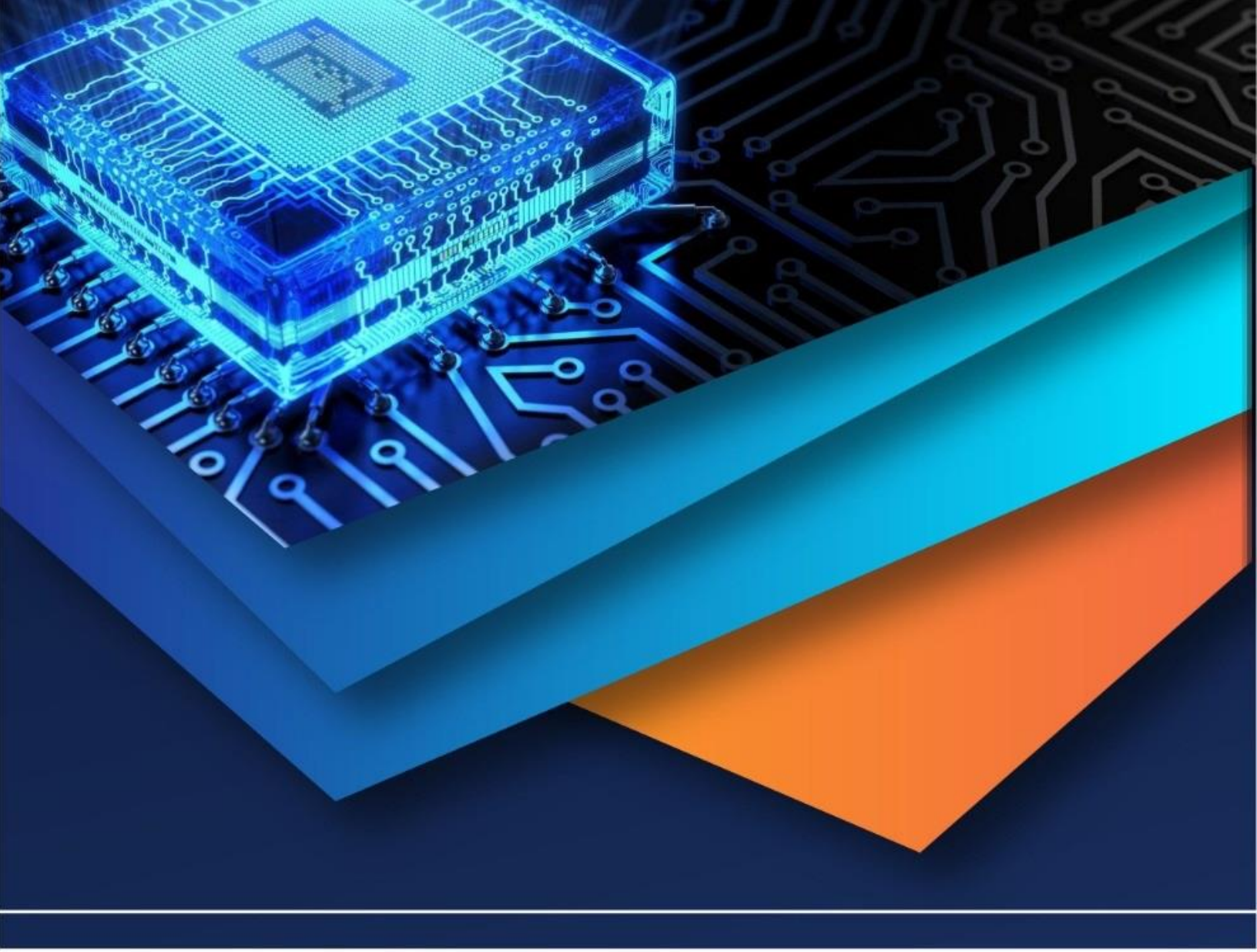
Fig 2:Feeding fighting recognition training data into model

## VI. CONCLUSION

Due to increase in thefts and crime in many areas today, one can require a system that can recognize the incident on real-time basis, with proper storage of evidence of the incident which can perform the task of sending alerts to nearby authorities. During the implementation, we focused on three aspects: Product should be Budget friendly, Must contain improved theft recognition models and should generate a quick alerts/notifications. By conducting the implementation for improving theft recognition. As we had used CNN with real time human gesture detection, we had implemented it successfully using MediePipe with accuracy of 75%. Also we had used a human hand gesture recognitionmodel using python libraries to detect the fist while beating that can generate the alerts. As it is easy and efficient technique to implement because according to our research there are only few such references implemented before in this domain.

## REFERENCES

[1] Gayashan Kariyawasam," Suspicious Activity Detection in Surveillance Footage", ICECTA – 2019.
[2] Garima Mathur, "Research on Intelligent Video Surveillance Techniques for Suspicious Activity Detection – Critical Review", ICRAIE- 2016.
[3] Cheng-Jia W an g , —" real-time car detection and driving safety alarm system with google TensorFlow object detection API", 2019.
[4] Nagarjuna R Vatti, —" Smart Road Accident_Detection_and_Communication System", IEEE-2018.
[5] Kanchana. V, —" Multiple Car Detection, Recognition andTracking in Traffic", INCET2020.
[6] Rasha Saffarini, "Smart System toavoid car accidents", ICPET-2020.
[7] Rajib Debnath, —" Automatic Visual Gun Detection Carried byA Moving Person", ICIIS-IEEE 2020.
[8] Sorina Smeureanu, —" Real-Time Deep Learning Method for Abandoned Luggage Detection in Video", EUSIPCO 2018.
[9] Website "Agile Model" (Software Engineering) – javatpoint,
[10]Website "TensorFlow" – Wikipedia.
[11]Website "Abandoned luggage detection using a finite stateautomaton in surveillance video"Requested PDF ,researchgate .net.
[12]Website "Convolutional Neural Network" – javatpoint.
[13]Website "Illumination-Sensitive Background Modelling Approach for Accurate Moving Object Detection", researchgate.net.
[14]Website "Object detection", Wikipedia.org.
[15]Website "Background subtraction – OpenCV",Geeks for Geeks.
[16]Website "what-is-OpenCV", www.tutorialspoint.com.
[17]Website "Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network", researchgate.net.
[18]Website "Number plate detection and OCR", www.youtube.com.js
[19]Website "OpenCV for suspicious object detection by Xvidia Technologies", xvidia.net.
[20]Ishan Dixit, "Debunking Convolutional Neural Networks (CNN) with practical examples, Becoming Human: Artificial Intelligence Magazine".
[21]Jing-Hao Sun, "Research on the Hand Gesture Recognition Based on Deep Learning".
[22]Kapitanov Alexander and team, "HaGRID — Hand Gesture Recognition Image Dataset".
[23]Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, "MediaPipe Hands: On-device Real-time Hand Tracking".
[24]Website "Hand Tracking and Gesture Recognition Using AI: Applications and Challenges"–Intellias.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)