DETECTING UNUSUAL CROWD BEHAVIOUR WITH OPENCV AND DEEP LEARNING

Shreya G Bhat*1, Disha N Deshmukh*2, Kavyashree N*3, Mrs. Meghashri E M*4

*1,2,3Department Of Computer Science And Engineering, Sai Vidya Institute Of Technology, Bengaluru, Karnataka, India.

*4Assistant Professor, Department Of Computer Science And Engineering, Sai Vidya Institute Of Technology, Bengaluru, Karnataka, India.

ABSTRACT

Suspicious behavior in open areas is sensitive and prone to serious consequences. There are many systems built based on video frame acquisition that monitors motion or identifies pedestrian but those systems are not smart enough to find out the unusual activities even in a real-life scenario. It is necessary to identify fugitive scenarios in real-time from video surveillance for rapid and immediate control before any casualties. To create a technology that can automatically identify suspicious behavior using computer vision, the system focuses on distinguishing suspicious scenarios and pinpointing the precise target of the activity. The system utilizes the OpenCV library for distinguishing and categorizing several kinds of tasks or actions in real-life scenario. Automated event identification, movement-based recognition, person count tracking, autonomous robot navigation, and a variety of other disciplines are some of the other topics covered. The identification of objects in movement and tracking movements from videos is very critical. Separating objects from their background, on the other hand, is a challenging process. So, it is very crucial to get clarity on the contents of the visual data and the background of the objects. A difficulty arises when an object appears from the background as a result of understanding the footage and its elements with portrayed scenarios becomes the most essential requirement. The utilization of a typical human behavior approach is the predictable goal in the unpredicted activity detection phase. The technique is created at the start while working with certain common activity datasets. Actual information and facts are compared with the pattern during the verification step. Finally, a judgment is made about whether the behavior is expected or not. Unusual activity detection in the real-world security system is challenging due to the requirement of a specified typical human activity approach.

Keywords: Crowd Activity, YOLO Algorithm Implementation, Motion Detection, Alerting Systems, Live Telecasting, Deep Learning Model, Video Surveillance, Clustering Mechanism, Social Distancing Measures.

I. INTRODUCTION

In general terms, it is shown that humans appear to be prone to many immune diseases which mechanisms to both genetic and environmental related diseases. Human health is also affected by various agents present in the surrounding environments which may be physical, social, or mental well-being, not just by the mere illness that the human body is affected with.

The virus responsible for this is the “Corona Virus” which is one of the RNA Viruses that target mammals and birds. Corona viruses are one among the family of viruses that infect human beings to certain illnesses that have preliminary symptoms such as acute respiratory symptoms, common flu, and few others.

The common public may adopt to maintain social distancing in certain public gatherings. Precautionary measures such as; staying indoors, limiting travel, avoiding crowded areas such as shopping centers, etc. can be implemented. Necessary unavoidable social events can be attended provided there are no contact greetings, distancing themselves from others of about 6-20 feet, and by wearing masks accordingly. Physical distancing helps limit the spread of COVID-19 or any similar sort of infectious disease. Acic Crowd detection allows for estimating the number of people in a defined area in many situations (cultural events, conferences, public sector areas like malls, etc.) Some of the top currently researched algorithms are:

a) Region-based Convolution Neural Networks (R-CNN)
b) Faster R-CNN
c) Spatial Pyramid Pooling (SPP-net)
d) YOLO (You Only Look Once)
e) Histogram of Oriented Gradients (HOG)
f) Region-based Fully Convolution Network (R-FCN)
g) Single Shot Detector (SSD)

With the tremendous advancements in the technological sector of non-expensive and excellent quality video recording equipment, modern computers have escalated their demand for analysis for such footage in every industry. Detecting moving objects and tracking them from the footage, on the other hand, is highly crucial and critical. Distinguishing objects from their surroundings is likewise a challenging job. Visual/Optical observation is one of the exceptionally recognized investigates, which has a more noteworthy region, human behavior application movement checking, public wellbeing in spaces like banks, shopping areas, private regions, and so forth. In a number of fields, motion-based recognition, human counting, autonomous robot navigation, and other types of automated event detection are implemented. CNN’s, the image configuration parameter for performing operations and neural networks (NNs) that integrate the image features from the CNNs with the picture that make up the system. This project is efficiently possible to demonstrate, with the use of Alerting Systems. The maximum probability determines the classification.

II. METHODOLOGY

There are four steps to constructing a deep learning model:

1) Gather your Data
2) Split your Data
3) Train Your Network
4) Evaluate

The first step in creating a deep learning network is to collect our initial dataset. Then require both of the images and the labels associated with each image. These labels must be drawn from a specific finite category. Now, we need to divide our initial data sample as two separate datasets:

a) Training Set
b) Testing Set

We can now train our network using the visuals from our training set. Our network’s objective here would be to learn how to classify each of the categories using the labeled data. When the model makes a mistake, it learns from this and seeks to improve. Evaluation of our trained network, we present each of the photos in our testing set to the network and asks it to predict what the label of the picture represents. The model’s predictions for each image in the testing set are then tabulated.

A. Working with the Data

In real-world applications, there is a need for bounding box detections in advance; therefore a detector is with the tracker. Assume we consider bounding box information for all items within a frame and then allocate IDs when provided with the information of the box for an ID in frame 1 using:

- Centroid based ID assignment:

  We can allocate IDs based on the centroid of the bounding boxes. In frame 1, we achieve this by computing the centroid for each bounding box. In frame 2, we evaluate the new centroid and assign IDs depending on their relative distance from the former centroid. The underlying principle is that the centroid shifts only slightly from frame to frame. This straightforward method works effectively as long as the centroids are separated by a sufficient distance. When persons are near to one another, this method fails because IDs may be switched.

- Kalman Filter:

Fig 1: Examples of common training and testing
Kalman Filter is an advanced approach over simple centroid-based tracking. The Kalman Filter allows tracking model based on an object's location and velocity and forecast where it is most likely to occur. It forecasts future location and velocity using Gaussian distributions. When it receives a new reading, it may use probability to assign the data to its predictions and updates itself. It has a small memory footprint and operates quickly. Thus it outperforms centroid-based tracking in terms of accuracy since it considers both location and velocity of motion.

B. Group Detection Implementation

- Detecting individuals in a video feed with the YOLO object detector.
- Determining the centroid for each detected person.
- Computing the pair wise distances between all centroids.
- Checking to see if any pair wise distances were < N pixels apart and if so, indicating that the pair of people violating.

Fig 2: Object Detection from the given point of location

C. Output of proposed model

Fig 3: GUI of the project

Fig 4: Object Detection of a person when not in crowd

III. ARCHITECTURE & ANALYSIS

Data Collection:
The camera and software in use take the real-time data on spot.

Relevant Dataset:
Since our work is regarding object detection, so we extracted. The existing training data set that we refer to train our model is extracted from any of the websites like Kaggle, or any other web scraping interfaces. We consider object that are actually in the surrounding of the camera area, else no detection of objects or Humans.

Data Preprocessing:
The data is highly unstructured, so it needs to be refined such that system can efficiently identify/recognize the change in patterns and notify the same. Various steps are being followed in this phase; the camera image is captured initially. If given permissions to store, then the captured images (data frames) are stored on to the memory after being processed. These processed data frames are the digital signals that are stored on the memory device/appliance for future validation. As much as possible we try to clean the noise or any unnecessary elements within the recorded frames. Labeling the model when any event occurs and retrieving
the existing data on the existing memory are also a part of the Data Preprocessing. Hence this stage ensures the classification with better accuracy.

**Future Engineering:**

From the preprocessed stage, various features are extracted as per the semantics and are converted into user understandable formats (Example: Buzzer alert – i.e. conversion of digital signals to sound signals. GPS Readings – Accurate tracking, Displaying of the footage is also a format that user understand via an output device).

![Proposed System Architecture](image)

**Figure 5:** Proposed System Architecture.

### IV. CONCLUSION

The configurations presented thus far are meant to comprehend fundamental human activity such as walking, running, and most other activities, but they are unsuitable for use in a busy space. The proposed device is capable of detecting irregular human behavior from the crowd and behavior using the motion effect map and OpenCV. The proposed system's accuracy rating is marginally higher than that of others, and the basic tracking algorithms play a pivotal role for any dealing with Previous Crime Appraisal. By using OpenCV and Motion Impact Diagram, the proposed device is capable of accurately detecting the irregular human behavior of the crowd, which improves the device's accuracy and expertise to a great degree. The availability of using the already existing GPS systems in our project demonstrates that we are trying to showcase the utmost a verified, accurate, and précised data currently in use. We will be able to use this Project for the real-time scenarios for COVID-19 norms.

### V. REFERENCES


