Violence Detection in Public Places using Machine Learning

[1]Priyam Bhatt, [2]Palav Desai, [3]Hina Rana, [4]Ishita Theba

[1]Student, [2]Student, [3]Student, [4]Assistant Professor,

Computer Engineering Department,

A. D. Patel Institute of Technology, Anand, India.

Abstract : Day by day modern technology is advancing and helping people to make their life easier. In recent years a large amount of surveillance cameras has been installed at various public places like traffic signals, offices, schools and so on. They are also being used during various public events like concerts, marriages, rallies, etc. The goal of this research article is to develop a violence detection system using these pre-installed surveillance cameras. We have put forward a system that detects violence from these videos and alerts the authorised person when violence is detected so that it can be handled and managed timely, causing the least harm to the public. The concept of machine learning is utilized here.

Keywords **- Machine Learning, Surveillance Cameras, Violence Detection, Image Processing, OpenCV, unusual activity detection, Convolutional Neural Networks.**

I. INTRODUCTION

India, being a developing country, encounters many day-to-day problems affecting the life of civilians to a great extent. As a result of rapid population growth, streets, gardens, malls and other public places are getting overcrowded. This will cause mass rush, riots and other violent offenses, which ultimately results in total loss of control.

Violence often occurs at large gatherings, events and other public places like concerts, fairs, stadiums, rallies and so on. In 2018 there were 76,851 rioting-related offenses across India, which when adjusted for a population, works out to be 5.7 riot offenses per 100,000 people.[1] At some public places, police and security guards are appointed to take care of such violent situations but they are not enough. There are high chances of mass panic, fights and accidents which require continuous surveillance and highly efficient violence detection. This type of violence prevention is not possible by traditional methods.

The main focus of the proposed system is on monitoring such violent incidents and giving an early intimation to the authorised person with minimum human intervention. Using Machine Learning, the system is smart enough to detect violence properly and handle such affairs efficiently.

II. WORKFLOW

2.1 Overview

The process starts with video capturing through various pre-installed surveillance cameras at traffic signals, during public gatherings, concerts, rallies and so on. Then it is passed through the pre-trained system where the video is checked frame by frame. If any kind of violence is detected then the authorised person is notified with an auto-generated alert text message.

2.2 Dataset

To build a good detection classifier, OpenCV needs to be trained on hundreds of images and videos containing scenes of some sort of violence. The training images and videos should have random scenes along with the desired violent and non-violent scenes, having a huge variety of background and lighting conditions. There should be images and videos with various kind of violence that is only halfway in the picture, partially obscured or overlapping with something else. The team has not only taken data from Kaggle but has also gathered data manually by converting online videos containing some sort of violence into images. There were almost 3000 images and videos with various violent scenes. All of the datasets are then labelled which eventually helps the detector in detecting violence.

2.3 Choosing OpenCV

To run Convolutional Neural Network (CNN) and other neural network-based computer vision architectures, OpenCV is widely used. It has an image processing module that includes geometrical image transformations (resize, affine and perspective warping, generic table-based remapping), linear and non-linear image filtering, colour space conversion, and so on.[2] It also has a video analysis module that includes motion estimation, background subtraction, and object tracking algorithms. [2]

2.4 System Architecture

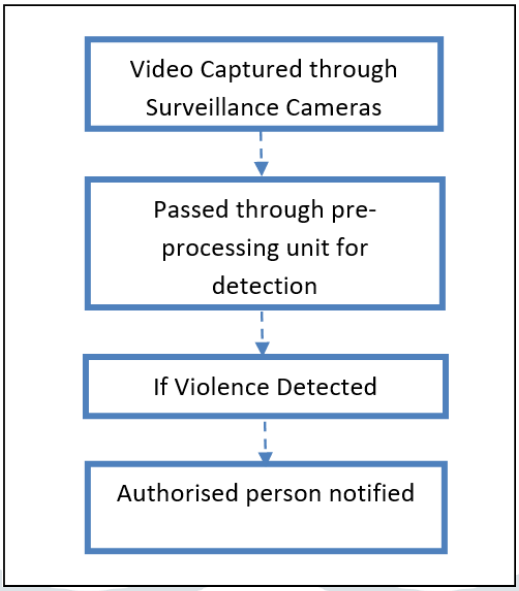

Fig.1 Violence Detection through videos using Machine Learning

2.5 Training and Testing

The available dataset is typically split into three sets: training, validating and testing. Loss values are calculated via forward propagation by a network trained using a training set. With the help of backpropagation, learnable parameters are updated. Wherever the concept of loss and the algorithm for optimizing gradient descent plays an important role, the backpropagation algorithm approach is broadly used to train neural networks. During the training cycle, a validating collection of data set is used to track the model performance. A test set is only used once at the very end to assess the efficiency of the final model. [3] [4]

For training, we have taken various videos and images of people doing normal activities in public places. After the collection of various videos and images, we placed them in a folder. We have taken video footage of a fight for testing the data. Optical flow is computed for each frame in the video. A frame has unique vectors of motion effect in it if a suspicious behavior of a person takes place, together with its adjacent block. [3] [5] We have used the farneback algorithm in the system as it processes all the pixels of the video. Farneback algorithm is a dense technique used for frame processing [6].

We then create a minimum distance matrix over the mega blocks after extracting the spatio-temporal element vectors for all mega blocks. In this corresponding mega block, the value of an element is determined by the minimum Euclidean distance between the mega block centre values and the element vector of the current test frame. The smaller the value of an element the less likely the block will have a violent activity. On the other hand, if there is a higher value then it can assume that in consecutive frames violence occur. Therefore, in the minimum-distance matrix, we consider the highest value to be the frame representative element value. If the maximum value of the minimum distance matrix is greater than the threshold then the frame is marked as violent. [3] [7]

For the testing of the notification system, we have used Fast2SMS.[8] A free API of 50 credits is provided by them for testing purposes. The API has been integrated with our python code of violence detection, so that whenever violence occurs the notification is received by authorised person.

3.6 Results

Once the videos are pre-processed, the next step is to detect violence. The videos will be examined frame by frame and the system will automatically try to detect violence in each frame, as shown in Figure 2. Once the violence has been detected, the authorised person will get an alert message that violence has been detected, as shown in Figure 3.

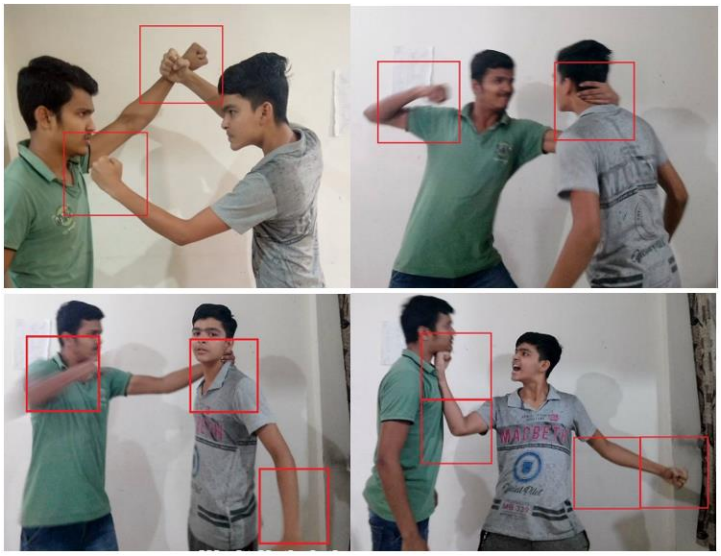

Fig 2 Violence Detected in each Frame

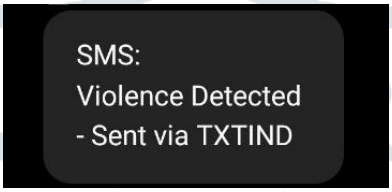

Fig 3 SMS received by authorized person

III. CONSTRAINTS AND CHALLENGES

The main task of this project is to detect violence in the video captured through surveillance cameras located at various places and notify the authorised person. Keeping this as primary focus there are some issues that can be considered as hurdles and need to be addressed carefully.

There will always exist some zones in cities where any kind of surveillance is prohibited. There is also a possibility that the video quality is not good enough for detection. Loss of connection and hacking of surveillance systems can also be considered as a major challenge. These all challenges need to be addressed properly following the security protocols.

IV. CONCLUSION

This proposed system is not only helpful to the security officials but also to event management teams. It helps in the detection of violence and notifying the authorised person at the proper time through text SMS like “Violence has been Detected”, so that any causalities that can occur are prevented.

REFERENCES

[1] https://www.livemint.com/news/india/the-geography-of-riots-in-india-11583304754521.html

[2] https://docs.opencv.org/master/d1/dfb/intro.html

[3] D. Lee, H. Suk, S. Park and S. Lee, "Motion Influence Map for Unusual Human Activity Detection and Localization in Crowded Scenes," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 10, pp. 1612-1623, Oct. 2015, doi: 10.1109/TCSVT.2015.2395752

[4] https://towardsdatascience.com/train-validation-and-test-sets72cb40cba9e7

[5] Li, Weixin, Vijay Mahadevan, and Nuno Vasconcelos. "Anomaly detection and localization in crowded scenes." IEEE transactions on pattern analysis and machine intelligence 36.1 (2013): 18-32.

[6] Farnebäck, Gunnar. "Two-frame motion estimation based on polynomial expansion." Scandinavian conference on Image analysis. Springer, Berlin, Heidelberg, 2003.

[7] Lucky, Bhupendra Panchal,” Unusual Crowd Activity Detection using Euclidean Distance in OpenCV for Creating Motion Influence Map”, in International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES), volume 5, Issue 07, July-2019

[8] https://docs.fast2sms.com/