



RAJIV GANDHI

INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

A PROJECT PROPOSAL ON “Violence Detection Using Surveillance Camera”

SUBMITTED BY

Mr. HARSHA R(1RG19CS017)	+91 6366 102 695 harsharharry@gmail.com
Mr. NITYANAND KUMAR(1RG19CS039)	+91 80517 16409 nityanandchaudhary420@gmail.com
Mr. SUBHASH B S(1RG19CS061)	+91 6364 446 418 subhashbs36@gmail.com
Mr. SUMANTH HEGDE(1RG19CS065)	+91 9483633976 sumanthdh13@gmail.com

UNDER THE GUIDANCE OF

Mrs. BHAGYASHRI WAKDE, Assistant Professor	+91 7026383816 bhagyashelke2015@gmail.com
Dr. ANTHONY RAJ, Assistant Professor	+91 9611096668 anthonyraj171@gmail.com

DATE OF COMMENCEMENT OF THE PROJECT:

22-August-2022

PROBABLE DATE OF COMPLETION OF THE PROJECT:

April 2023

Guide
Mrs. Bhagyashri Wakde

HOD
Mrs. Arudra A

Violence Detection using Surveillance Camera

Abstract:

Movenet Pose Estimation is a technique that uses deep learning algorithms to detect human body posture and movements in real-time. In the context of violence detection using surveillance cameras, Movenet Pose Estimation can be utilized to analyze the video stream and detect any abnormal or violent behavior. The system outputs a set of keypoints representing the position of joints and limbs, which can then be used to identify specific movements and postures that are indicative of violence. This allows for real-time detection and response to violent incidents, improving the overall security of the monitored location. Movenet Pose Estimation can also be customized to recognize specific movements and postures in different environments and integrated with other security systems such as alarms and access control. However, limitations such as the need for high-quality video footage and robust algorithms must be considered. . it can also detect fire, car crash, violence in office, people walking in Street, car parking area, fight in Street so on..... Finally, it will alert the system through Telegram using Telegram-Bot.

Introduction:

Video surveillance has become a common tool for ensuring security in both public and private spaces. With advancements in digital video recording, network video transmission, and video analytics, real-time monitoring and analysis of activities has become easier and more cost-effective. However, manual monitoring of the massive amount of video footage generated by these systems can be time-consuming, resource-intensive, and subject to human error. This has led to the development of computer vision and deep learning algorithms that automate the monitoring process and improve the efficiency and accuracy of video surveillance systems.

One such technology is Movenet Pose Estimation, a deep learning-based technique that detects human body posture and movements in real-time. This technology provides a fast and effective way to detect and respond to violent incidents and is gaining popularity as a tool for violence detection using surveillance cameras. This report will delve into the applications of Movenet Pose Estimation for violence detection, including an overview of how the system works, its benefits and limitations, and ways in which it can be integrated with other security systems. Additionally, privacy concerns related to the use of video surveillance will be addressed and the cost-effectiveness of Movenet Pose Estimation as a violence detection system will be discussed.

In the context of violence detection, Movenet Pose Estimation analyzes video streams from surveillance cameras to identify any abnormal or violent behavior. The system outputs a set of keypoints that represent the position of joints and limbs, which can then be used to detect specific movements and postures associated with violence. This enables real-time detection and response to violent incidents, enhancing the overall security of the monitored area. Movenet Pose Estimation can be customized to recognize movements and postures specific to different environments and can be integrated with other security systems such as alarms and access control.

The deep learning algorithms used by Movenet Pose Estimation have been trained on large datasets of human body movements and postures. These algorithms are robust to variations in lighting, camera angle, and body pose, making Movenet Pose Estimation a reliable tool for violence detection. Additionally, the use of Movenet Pose Estimation can provide a cost-effective alternative to traditional security measures by reducing the need for human security personnel and providing real-time alerts and responses to violent incidents.

However, the use of video surveillance systems, including Movenet Pose Estimation, raises privacy concerns. The collection and storage of video footage of individuals raises questions about the privacy of personal information and the potential for abuse. It is important to consider the ethical and legal implications of video surveillance and ensure that privacy rights are protected.

In conclusion, Movenet Pose Estimation is a valuable tool for violence detection in video surveillance systems. Its ability to automate the monitoring process and detect violent incidents in real-time can enhance the security of monitored areas and provide a cost-effective alternative to traditional security measures. While privacy concerns must be taken into consideration, the use of Movenet Pose Estimation can be an effective means of promoting safety and security. Nayan Nirikshan project also detects the some of the outdoor and indoor cases like car crash, people walking on a street, car parking area, fight on a street, fire on a street, office, buildings etc... Finally, it will alert the system through Telegram using Telegram-Bot.

Motivation

The motivation for a violence detection using surveillance cameras project comes from the need to improve public safety and respond more quickly to violent incidents. In recent years, incidents of violence have become increasingly common, and traditional methods of detecting and responding to violence, such as eyewitness reports or manual video analysis, are often slow and unreliable. By using AI and computer vision techniques to automatically detect instances of violence, it is possible to improve the speed and accuracy of violence detection, and ensure a quicker response time by relevant authorities, such as law enforcement.

Additionally, the use of surveillance cameras has become increasingly common in public spaces, and there is a growing need to make the most of this data to improve public safety. By developing an AI-based violence detection system, it is possible to leverage the wealth of data generated by surveillance cameras to make our communities safer and more secure.

Problem Definition

The problem definition for a violence detection using surveillance cameras project is to develop an AI-based system that can automatically detect instances of violence in real-time, using video footage from surveillance cameras. The system should be able to accurately detect a wide range of violent behaviors, including physical altercations, weapon usage, and other forms of aggression, and provide timely alerts to relevant authorities, such as law enforcement.

Proposed System:

The image from the video is passed to yolo v3 where each person is detected and each frame is passed with its bounding box coordinates to move net lightning where their key point is detected and passed to a black image where it is plotted.

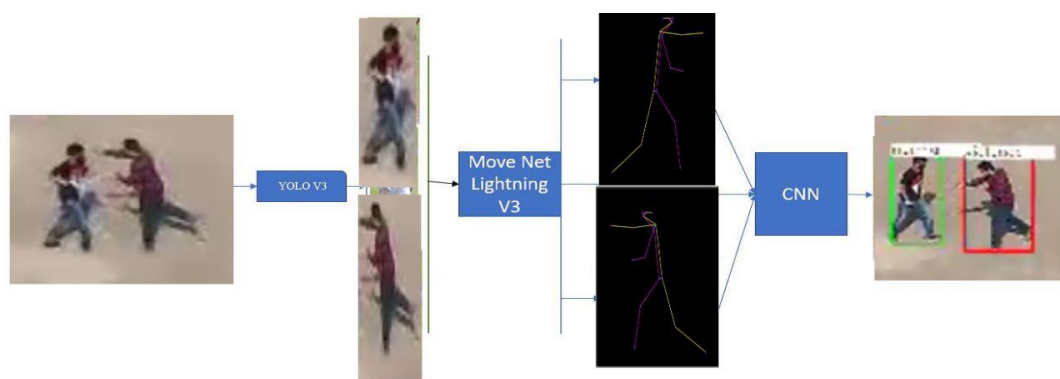


Fig:Proposed System

Once obtained the pre-processed image it is rescaled to 180x180, normalized and later passed to CNN where it provides us with a result of Violence or non-Violence. These results are passed for each frame with their bounding box with the original frame. Only the bounding box which meets the criteria for the Violence action is considered for display. This result is later printed onto the bounding box having Violence action w.r.t original image. Each frame is processed individually one after another which is helpful for us to obtain ROI (region of interest). As seen in figure 3 each person is detected and passed to move net individually as for multi-pose detection the results are shown that in move net their key-points are not accurate for the body pose hence we pass each frame individually through single-pose detection.

System Architecture:

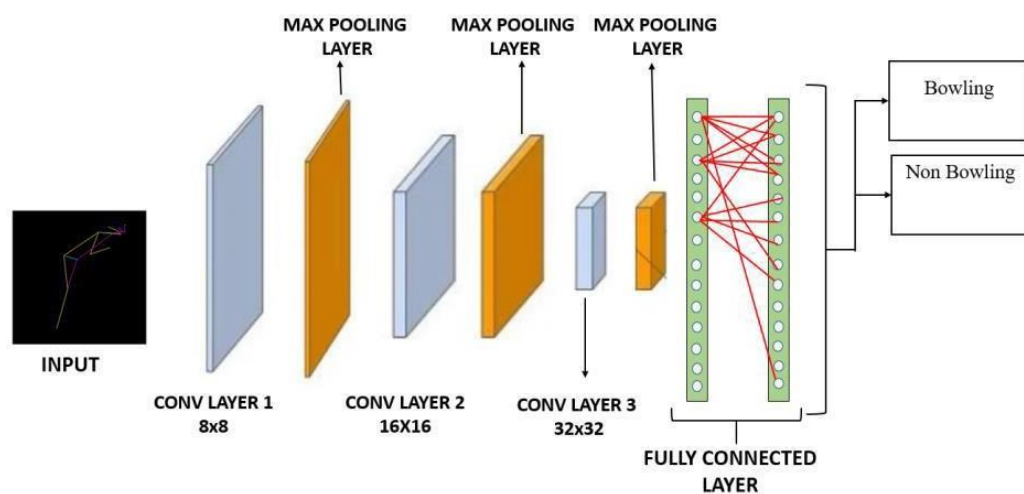


Fig:CNN architecture

The Convolution Neural Network (CNN) is based on the Rectifier Linear Unit (ReLU) as its activation function found to have time efficiency for both training and testing. ReLU is a type of activation function which yields the input as output if the input is positive, else it results in a null value. The input image is resized to 180x180 for the processing in CNN these resizing of images are done.

The CNN is designed to have 3 convolution layers along with 3 max pooling layers and one fully connected or dense layer. There are 2 output classifications bowling and non-bowling action. The image generated by drawing the key points on a black image is the input to the CNN. The CNN is used for classifying the action is of bowling or of non-bowling.

If classified as bowling with a confidence level of the model is less than 65% is considered as non-bowling action this is done to reduce the false positive. Total number of parameters present in the network are 992,146 out of which the first convolution layer is of the size 8x8 (conv layer 1) convolution filter followed by size 16x16 (conv layer 2) and 32x32 (conv layer 3). Parameters that are present in the dense layer are 991296. The architecture of the convolutional neural network that is in use is as mentioned in Fig.

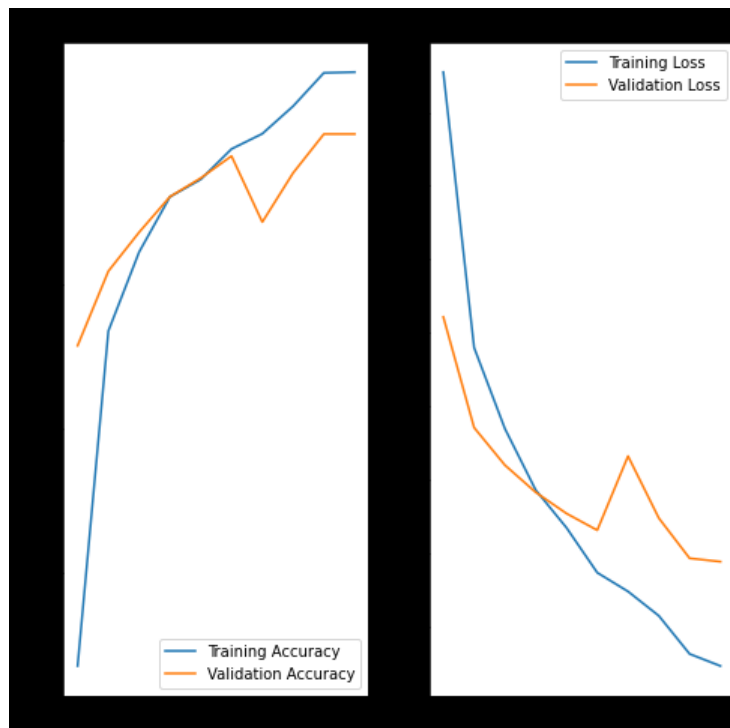


Fig: Model accuracy and loss during training

Statement of Objectives:

- To develop an AI-based system that can automatically detect instances of violence in real-time using video footage from surveillance cameras.
- To accurately distinguish violent incidents from other types of behavior and classify instances of violence into different categories, such as physical altercations, weapon usage, and other forms of aggression.
- To provide real-time alerts to relevant authorities, such as law enforcement, as soon as an instance of violence is detected.
- To minimize privacy and civil liberty concerns associated with the use of cameras for surveillance, while still achieving the goals of the project.
- To develop a scalable and deployable system that can handle large amounts of video footage from multiple cameras and be easily deployed in a variety of environments.
- To evaluate the performance of the system using metrics such as accuracy, precision, and recall.
- To provide recommendations for improving the system based on the results of the evaluation and to address any limitations or challenges that arise during the project.

Hardware and Software Requirements:

Hardware Requirements:

·System	:	i3 6gen or above.
·Hard Disk	:	250 GB.
·Monitor	:	15 VGA Color.
·Mouse	:	Logitech.
· RAM	:	8 GB
-Graphics card	:	RTX 1070 or above
-Surveillance Camera :		Wireless, Day/night, Network/IP CCTV cameras

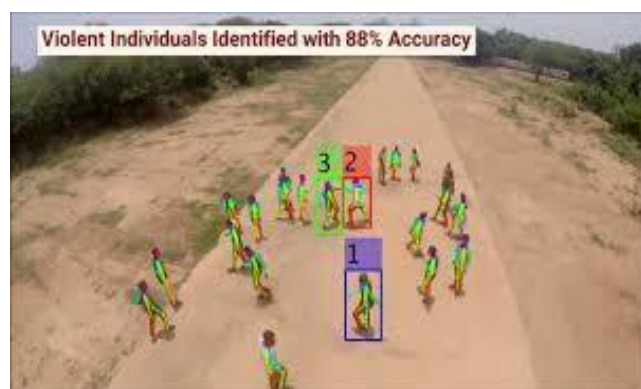
Software Requirements:

·Operating system	:	Windows / Linux.
·Coding Language	:	Python, Django, reactJs
. Software Tool	:	Movenet, OpenCV, Django, ReactJs
Internet Access		

Possible Output:

A violence detection using surveillance camera project can generate real-time alerts, captured video footage, violence reports, and improved public safety. The system can detect violent activities and send alerts for prompt response, and the video footage can be used for investigation and as evidence. Reports and statistics on the frequency and location of violent incidents can also be generated, allowing for targeted intervention strategies.

Conclusion and Future Enhancement:



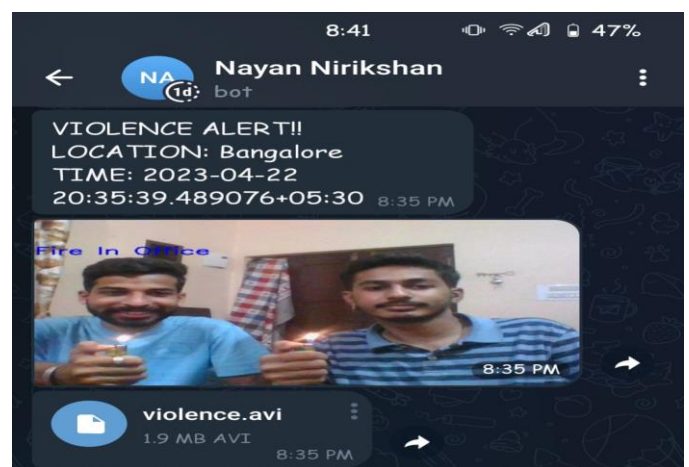
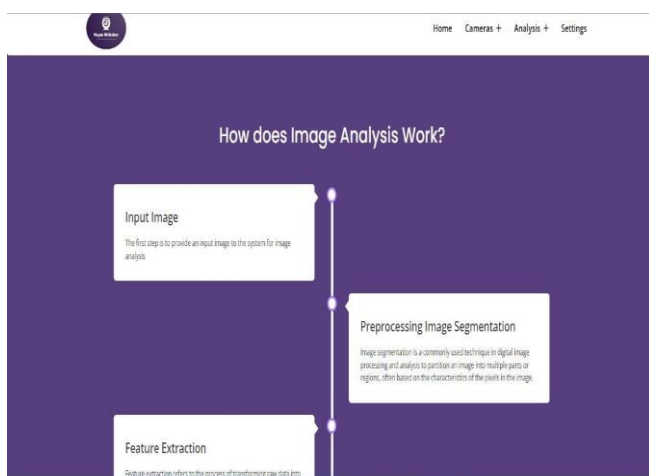
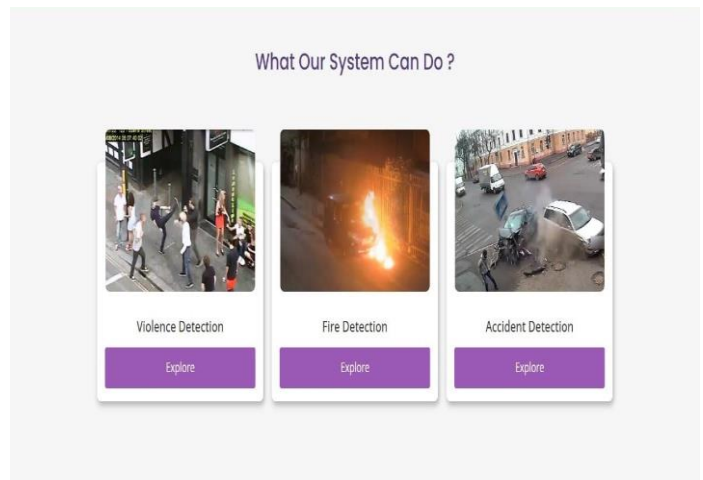
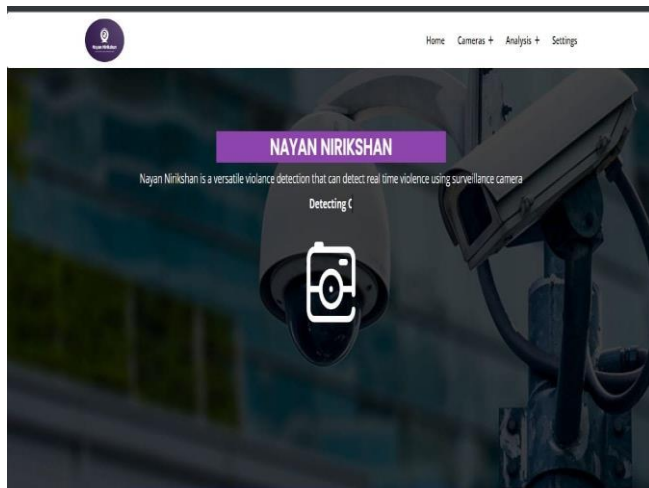


Fig : Snapshots

Conclusion : A violence detection using surveillance camera project is a promising solution for improving public safety by detecting and responding to violent activities. The integration of advanced computer vision, artificial intelligence, and machine learning algorithms enables the system to accurately detect violence in real-time. Future enhancements: In the future, the violence detection system can be further enhanced by incorporating multi-camera support, advanced deep learning algorithms, and real-time video analytics. The integration of additional sensors, such as audio sensors, can also improve the accuracy and reliability of the system. Furthermore, the use of blockchain technology can enhance the security and privacy of the system.

References:

- [1] Y. Wang, J. F. Doherty, and R. E. Van Dyck, "Moving object tracking in video," in Proceedings - Applied Imagery Pattern Recognition Workshop, 2000, vol. 2000-January, pp. 95–101.
- [2] B. Tian, Q. Yao, Y. Gu, K. Wang, and Y. Li, "Video processing techniques for traffic flow monitoring: A survey," in IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2011, pp. 1103–1108.
- [3] J. M. B. Oñate, D. J. M. Chipantasi, and N. D. R. V. Erazo, "Tracking objects using Artificial Neural Networks and wireless connection for robotics," J. Telecommun. Electron. Comput. Eng., vol. 9, no. 1–3, pp. 161–164, 2017.
- [4] M. Brown, J. Funke, S. Erlien, and J. C. Gerdes, "Safe driving envelopes for path tracking in autonomous vehicles," Control Eng. Pract., vol. 61, pp. 307–316, 2017.
- [5] V. A. Laurence, J. Y. Goh, and J. C. Gerdes, "Path-tracking for autonomous vehicles at the limit of friction," in Proceedings of the American Control Conference, 2017, pp. 5586–5591.
- [6] S. Walker et al., "Systems and methods for localizing, tracking and/or controlling medical instruments," 2017.
- [7] M. Buric, M. Pobar, and M. Ivacic-Kos, "An overview of action recognition in videos," 2017 40th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2017 - Proc., pp. 1098–1103, 2017.
- [8] P. Viola and M. Jones, "Managing work role performance: Challenges for twenty-first century organizations and their employees," Rapid Object Detect. using a Boost. Cascade Simple Features., 2001.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, 2005.
- [10] N. Wang, J. Shi, D. Yeung, J. J.-P. of the IEEE, and undefined 2015, "Understanding and diagnosing visual tracking systems," openaccess.thecvf.com.
- [11] P. Li, D. Wang, L. Wang, and H. Lu, "Deep visual tracking: Review and experimental comparison," Pattern Recognit., vol. 76, pp. 323–338, 2018.
- [12] Heidari and P. Aarabi, "Real-time object tracking on iPhone," in Lecture Notes