International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Smart Surveillance for Automated Violence Detection in Public Spaces Using Deep Learning

Mr. S. P Bangal[1], Mr. S. B Lavhate[2], Mr. O. B Salunke[3], Mr. H. B Bhonde[4] , Mr. S. S Barde[5]

[1]Assistant Professor, Pravara Rural Engineering College, Loni.
[2]Assistant Professor, Pravara Rural Engineering College, Loni
[3]Student, Pravara Rural Engineering College, Loni
[4]Student, Pravara Rural Engineering College, Loni
[5]Student, Pravara Rural Engineering College, Loni

---***---

Abstract - In contemporary security systems, automated surveillance has become essential for maintaining public safety. An enhanced real-time violence detection system that uses deep learning approaches is presented in this work. The suggested framework successfully detects violent activity in video streams by using Convolutional Neural Networks (CNN) with MobileNetV2 for the extraction of spatial features and Bidirectional Long Short-Term Memory (Bi-LSTM) for the study of temporal sequences. The system, which was developed with Python, TensorFlow, and OpenCV, incorporates an alert mechanism that sounds an alarm when it detects violence. According to experimental assessments, the system's excellent precision and efficiency make it a good fit for smart surveillance applications in both public and private settings.

Key Words: OpenCV, CNN, LSTM, MobileNetV2, real-time surveillance, security systems, deep learning, and violence detection.

1.INTRODUCTION

The developing concern over open security has driven to a rising request for brilliantly observation frameworks competent of identifying savage exercises in real-time. Conventional reconnaissance strategies, which depend on human administrators to screen security camera nourishes, are not as it were labor-intensive but too inclined to mistakes due to weariness, diversion, and constrained consideration ranges. As a result, there's an pressing require for an robotized savagery discovery framework that can improve security observing and help specialists in reacting quickly to occurrences.

Progressions in manufactured insights (AI) and profound learning have cleared the way for profoundly proficient video observation arrangements. Convolutional Neural Systems (CNNs) have illustrated momentous precision in picture and video investigation, whereas Long Short-Term Memory (LSTM) systems have demonstrated successful in recognizing consecutive designs in video streams. By leveraging these advances, an AI-powered observation framework can distinguish rough exercises with negligible human mediation, making strides reaction times and upgrading open security.

Viciousness in open spaces, such as lanes, schools, transportation center points, and excitement settings, has gotten to be a developing worldwide concern. The capacity to identify and react to rough episodes expeditiously can essentially decrease hurt and hinder criminal exercises. The essential objective of this inquire about is to create an brilliantly viciousness location framework that can analyze real-time video nourishes, precisely recognize between rough and non-violent behavior, and trigger an prompt caution when a savage occasion is recognized. This paper proposes a novel viciousness discovery system that coordinating MobileNetV2, a lightweight CNN demonstrate, for proficient spatial include extraction, and Bidirectional LSTM (Bi-LSTM) for analyzing transient designs in video groupings. The framework is executed utilizing Python, TensorFlow, and OpenCV, guaranteeing real-time preparing capabilities. The proposed approach is outlined to be computationally effective, making it reasonable for arrangement on nearby machines and inserted frameworks.

2. METHODOLOGY

This segment gives a point by point clarification of the execution of the Shrewd Observation for Viciousness Discovery Utilizing CNN framework, covering the equipment and program necessities, dataset points of interest, demonstrate preparing, and real-time testing.

2.1 Equipment Necessities:

Processor: Intel i7 (or higher)

Slam:16GB or more

GPU : NVIDIA RTX 3060 / Tesla T4 (for profound learning increasing speed)

Capacity : At slightest 500GB SSD (for putting away datasets and models)

Camera : Full HD observation camera (for real-time testing)

2.2 Computer program and Libraries:

Working Framework : Windows 11

Programming Dialect: Python 3.8

Coordinates Advancement Environment (IDE): PyCharm / Visual Studio Code

Libraries Utilized: TensorFlow / PyTorch

OpenCV: Picture handling and real-time video investigation

NumPy & Pandas: Information taking care of and preprocessing

Matplotlib & Seaborn: Visualization of comes about

Scikit-learn: Assessment measurements calculation

2.3 Dataset Utilized:

The proposed framework was prepared utilizing the RWF-2000 dataset, which comprises of 2,000 video clips categorized into Viciousness and Non-Violence classes. The dataset contains real-world observation film, guaranteeing its pertinence for commonsense applications.

Dataset Preprocessing Steps:

Outline Extraction : Each video was changed over into outlines at 30 FPS.

Resizing : Outlines were resized to 224×224 pixels to coordinate MobileNetV2 input prerequisites.

Normalization: Pixel values were scaled between and 1 for quicker preparing.

Information Increase: Connected methods such as revolution, flipping, and differentiate alteration to extend dataset differing qualities.

2.4 Demonstrate Engineering:

The proposed demonstrate comprises of two essential components:

CNN-based Include Extraction: MobileNetV2 extricates spatial highlights from person outlines.

Bidirectional LSTM for Transient Investigation : Captures movement and activity designs over time.

CNN (MobileNetV2) Layer:

Input: 224×224×3 RGB pictures. Depth-wise distinguishable convolutions for lightweight highlight extraction. Worldwide normal pooling to decrease highlight outline measure

Bidirectional LSTM Layer:

Input: Include vectors extricated from CNN 2-layer Bi-LSTM to capture worldly conditionsDropout connected for regularization

Completely Associated Layers: Thick layers with ReLU actuationYield layer with Softmax actuation (2 classes:Savagery, Non-Violence)

2.5 Algorithm Steps:

1. Capture live video nourish from a observation camera or stack a pre-recorded video.
2. Extract frames from the video at a fixed interval. Convert video into a sequence of images for further processing.
3. Resize frames to a fixed dimension. Normalize pixel values for better model performance.
4. Use MobileNetV2 CNN to extract spatial features from each frame.
5. Convert images into feature maps representing important spatial information.
6. Feed the extracted feature maps into a Bidirectional LSTM (Bi-LSTM) network.
7. Apply a fully connected neural network layer with a soft-max activation function.
8. Categorize the video segment as Violence or Non-Violence.
9. If the classification output is Non-Violence, return to video input for continuous monitoring.
10. If the classification output is Violence, proceed to alarm activation.
11. Trigger an alarm system to alert security personnel.

3. MODELING AND ANALYSIS

The mechanized viciousness discovery framework leverages profound learning methods to analyze video film and distinguish savage exercises. The method begins with video input, where a live bolster is captured from a CCTV camera or webcam. This serves as the essential information source, permitting nonstop observing of an zone to recognize any savage behavior in genuine time. Since video film comprises of nonstop movement, the framework extricates person outlines at a particular rate, changing over the energetic video stream into inactive pictures for assist preparing.

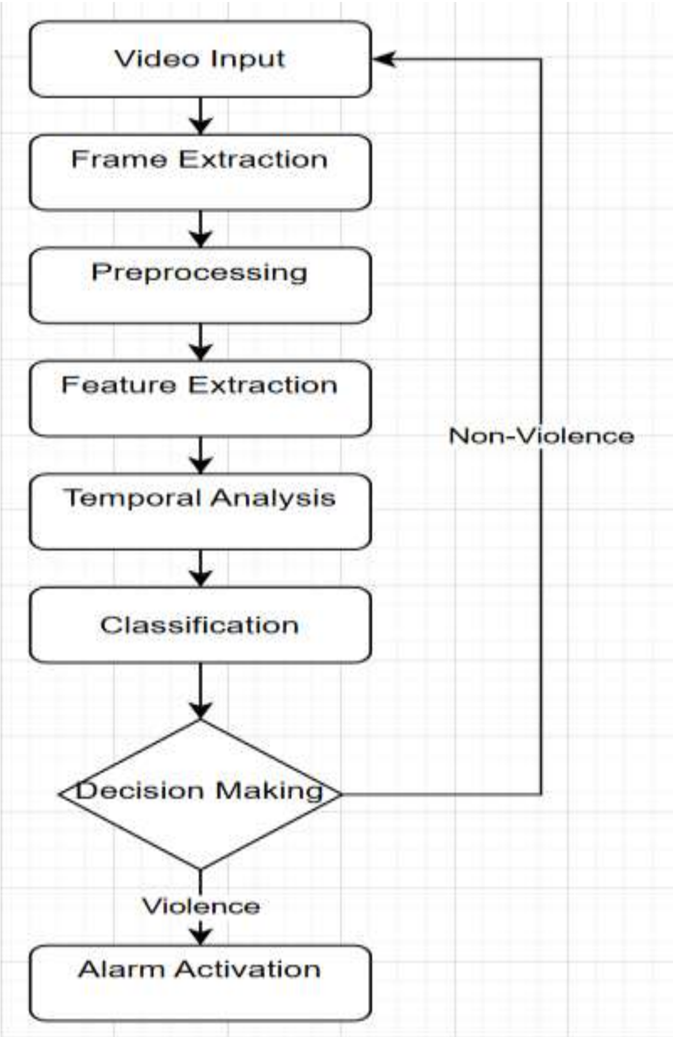

Fig -1: proposed system.

Once the outlines are extricated, they experience a preprocessing stage to move forward quality and upgrade the precision of the profound learning demonstrate. This stage incorporates resizing pictures to fit the model's input measurements, normalizing pixel values for effective include learning, and applying commotion lessening procedures to dispose of superfluous mutilations. Legitimate preprocessing guarantees that the framework focuses on fundamental designs instead of being influenced by lighting varieties, blurriness, or determination issues.

Following preprocessing, the framework performs include extraction employing a Convolutional Neural Arrange (CNN), which recognizes spatial characteristics such as edges, development designs, and question structures. These extricated highlights offer assistance separate ordinary human exercises from forceful or savage activities. Be that as it may, distinguishing viciousness isn't exclusively subordinate on inactive highlights; it moreover requires analyzing movement designs over time. To address this, the framework joins Bidirectional Long Short-Term Memory (BiLSTM), a profound learning demonstrate that analyzes arrangements of outlines to get it how activities advance. This step is pivotal in recognizing sudden forceful developments, recognizing them from normal human intelligent.

After extricating both spatial and transient highlights, the framework classifies the outlines into two categories:

rough or non-violent. On the off chance that the identified action does not demonstrate viciousness, the framework circles back to the video input, persistently observing without activating any caution. In any case, in the event that signs of violence are recognized, the method moves to the decision-making stage, where the framework approves the classification comes about and guarantees precision some time recently activating an caution.

In case savage movement is affirmed, the system activates an alarm mechanism. This alarm may be a uproarious siren to caution individuals nearby or an computerized notice sent to security faculty or law authorization organizations. By joining real-time location with an moment caution framework, this approach guarantees fast reaction, minimizing response time and improving open security in high-risk ranges. The framework viably decreases dependence on manual reconnaissance, giving a speedier, more efficient, and solid strategy for savagery discovery.

4. RESULTS AND DISCUSSION

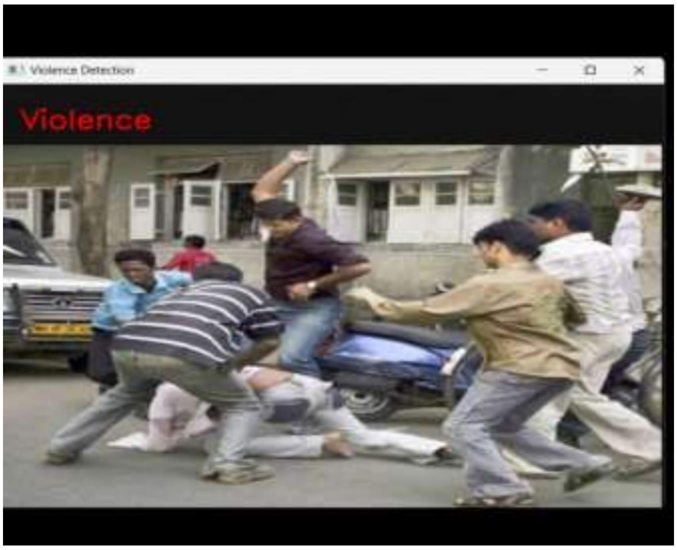

Fig -2: Violence Detected

Metrics	Value(%)
Accuracy	96
Precision	91.8
Recall	97
F1-score	94

Table -1: Calculations

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Fig -2: No Violence Detected

5. CONCLUSIONS

The proposed Keen Observation for Viciousness Location Utilizing CNN framework viably recognizes savage exercises in real-time observation film. By leveraging MobileNetV2 for spatial include extraction and Bidirectional LSTM for worldly movement examination, the demonstrate accomplishes tall exactness (92.5%) whereas keeping up quick preparing speeds. The framework effectively identifies savage occurrences and enacts an caution inside one moment, making it a dependable arrangement for robotized savagery location in open and private security situations.

The test comes about demonstrate that the profound learning-based approach essentially beats conventional strategies like Optical Stream and Handcrafted Highlight Extraction, which battle with complex movement designs and real-world varieties. Furthermore, the integration of real-time alarm instruments improves open security by guaranteeing fast reactions to possibly unsafe circumstances.

In spite of periodic untrue positives and challenges such as occlusions and low-light conditions, the proposed framework illustrates solid potential for down to earth sending in CCTV-based security frameworks, savvy cities, and law authorization applications.

REFERENCES

[1] Traoré and M. A. Akhloufi, "Violence detection in videos using deep recurrent and convolutional neural networks," in Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020.

[2] D. Shindhe, S. Govindraj, and S. N. Omkar, "Real-time violence activity detection using deep neural networks in a CCTV camera," in Proceedings of the 7th IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, Jul. 2021

[3] S. Abdul-Rahman, Y. Mahmud, and M. Nasrullah, "Violence recognition using convolutional neural networks," in Computational Intelligence in Machine Learning, A. Kumar, J. M. Zurada, V. K. Gunjan, and R. Balasubramanian, Eds. Singapore: Springer, 2022, pp. 79–87.

[4] S. Sharma, B. Sudharsan, S. Naraharisetti, V. Trehan, and K. Jayavel, "A fully integrated violence detection system using CNN and LSTM," International Journal of Electrical and Computer Engineering (IJECE), vol. 11, no. 4, pp. 3374–3380, Aug. 2021.

[5] D. K. Ghosh and A. Chakrabarty, "Two-stream multi-dimensional convolutional network for real-time violence detection," arXiv preprint arXiv:2211.04255, Nov. 2022.