BANGALORE INSTITUTE OF TECHNOLOGY

INFORMATION SCIENCE AND ENGINEERING

FINAL PROJECT (18CSP83) PRESENTATION
on

WatchfulEye : Recognition of Abnormal Human Activities in Wild Videos

TEAM MEMBERS :

HARSHA R B	1BI20IS035
HARSHITH V	1BI20IS036
KAVYA	1BI20IS044
MANAS C	1BI20IS050

GUIDE :

Mrs. Shilpa T
Assistant Professor
Dept. of Information Science and Engineering
Bangalore Institute of Technology

CONTENTS

ABSTRACT

- Existing systems for abnormal human activities recognition have some limitations such as not accurately detecting aggressive motions in dynamic and unstructured environment and also produce high false positive rate.
- To overcome these limitations we have developed a system called “WatchfulEye”.
- This project use Lucas-Kanade optical flow algorithm to precisely detect aggressive motions in a video.
- These aggressive motions are then processed using Canny Edge Detection to obtain critical edge features.
- Finally a custom CNN is developed and trained for robust recognition of abnormal human activities.

INTRODUCTION

- Video surveillance systems have long been used and evolved from analog to digital systems .
- Images or videos captured by a surveillance camera are stored for 3–15 days depending on the environment of the system and are renewed in a first-in, first-out order.
- The detection and analysis of anomalous activity in videos have become increasingly important in many real-time applications like passenger monitoring, and public safety.
- Differentiating between normal and abnormal human activities poses a significant challenge due to their inherent similarities and challenges like occlusion, clutter etc.
- We aim to develop a system which can identify anomalous actions from video sequences which have occluded scenarios in groups.
- It effectively differentiates between various normal and abnormal activities in the given input videos.

LITERATURE SURVEY

Sl No.	Title And Author	Year	Contributions	Drawbacks	Paper
1	A Hierarchical Spatio-Temporal Graph Convolutional Neural Network for Anomaly Detection in Videos Xianlin Zeng , Yalong Jiang, Wenrui Ding , Hongguang Li, Yafeng Hao.	2023	This paper proposes a Hierarchical Spatio-Temporal Graph Convolutional Neural Network (HSTGCNN) for anomaly detection in videos. It addresses existing model limitations by incorporating high-level graph representations for interactions among multiple identities and low-level graphs for individual body postures. This dual approach enhances the model's ability to detect anomalies more accurately.	The paper overlooks potential computational complexity and real-time applicability in dynamic environments, and relies on high-resolution video data, limiting effectiveness with lower-quality footage. Performance of the model in videos with occlusion and overlapping is not explored.	IEEE Transactions on circuits and systems for Video Technology, VOL. 33, No. 1, January 2023.
2	Motion Influence Map for Unusual Human Activity Detection and Localization in Crowded Scenes Dong-Gyu Lee, Heung-Il Suk, Sung-Kee Park, and	2023	This paper proposes a novel method for unusual human activity detection in crowded scenes. Specifically, rather than detecting or segmenting humans, it devised an efficient method, called a motion influence map, for representing human activities.	Motion-based methods might be sensitive to noise in the input data. If the method relies solely on motion information, it might not capture unusual activities that do not involve significant motion.	IEEE Transactions on Circuits And Systems For Video Technology.

Sl No.	Title And Author	Year	Contributions	Drawbacks	Paper
3	Abnormal Crowd Behavior Detection Using Motion Information Images and Convolutional Neural Networks CEM DIREKOGLU	2020	This paper introduces a novel method for abnormal crowd event detection where first, optical flow vectors are computed to generate a motion information image (MII) for each frame, and then MIIs are used to train a convolutional neural network (CNN) for abnormal crowd event detection.	The robustness of the MII approach to real-world challenges, such as varying lighting conditions, camera perspectives, and environmental disturbances, is not explicitly evaluated.	Journal
4	Transfer Learning based Video Surveillance for Abnormal Activity Detection Dr. R. Sittalatchoumy , Dr. S.Ewins Pon Pushpa, Karthick R, Rafiq Azharudheen M A, Dhiyanesh K V	2023	This paper focus to develop a novel system to automate the task of videotape surveillance and identify the unusual activities similar to road accident discovery, fall discovery, fire discovery, and weapon discovery.	This does not perform well in the crowded environments. Crowded environments may lead to occlusions, where objects may be partially or fully obscured by other objects or people. This could impact the system's ability.	7th International Conference on Intelligent Computing and Control Systems (ICICCS).

Sl No.	Title And Author	Year	Contributions	Drawbacks	Paper
5	Smart Video Surveillance Based Weapon Identification Using Yolov5 Dr. S. Nikkath Bushra, Ms. G. Shobana, K. Uma Maheswari, Nalini Subramanian	2022	This paper proposes a system which helps in identifying weapons held by a person as well as face recognition to identify the suspicious user using Yolov5, which is 88% faster than yolov4 in Deep Learning.	Weapon identification systems may be sensitive to variations in lighting and environmental conditions. The performance of the system might degrade in low-light situations or when there are other environmental challenges.	International Conference on Electronic Systems and Intelligent Computing (ICESIC)
6	SwinAnomaly: Real-Time Video Anomaly Detection Using Video Swin Transformer and SORT A. Bajgoti, R. Gupta, P. Balaji, R. Dwivedi, M. Siwach and D. Gupta	2022	This paper introduces a method , a video anomaly detection approach based on a conditional GAN-based autoencoder with feature extractors based on Swin Transformers. We utilize patch-wise mean squared error and Simple Online and Real-time Tracking (SORT) for real-time anomaly detection and tracking.	The proposed approach is the high sensitivity of Swin Transformers towards different lighting conditions, which can affect the performance of the model if it is not trained on different lighting conditions.	IEEE Access 2023

Sl No.	Title And Author	Year	Contributions	Drawbacks	Paper
7	Real-Time Anomaly Detection for Smart and Safe City Using Spatiotemporal Deep Learning Rabia Hasib, Atif Jan, Gul Muhammad Khan	2022	This paper presents an automatic recognition of unusual human behavior captured by a CCTV camera in public areas, using spatial temporal 3D convolutional neural networks.	This paper classifies only few categories like robbery, fighting, assault and etc.	2nd International Conference on Artificial Intelligence (ICAI).
8	Suspicious Human Activity Recognition using 2D Pose Estimation and Convolutional Neural Network Arjun S. Dileep, Nabilah S. , Sreeju S , Farhana K. and Surumy S.	2022	This paper proposes a real-time suspicious human activity recognition with high accuracy by introducing a Convolutional Neural Network and using the 2D pose estimation technique to the system.	This paper focus on only two suspicious activities, fall and wall climbing or trespassing.	International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)

Sl No.	Title And Author	Year	Contributions	Drawbacks	Paper
9	Revolutionizing Anomaly Detection in Surveillance Footage with ARLSTM Sanika Tanmay Ratnaparkhi, Aamani Tandasi, Shipra Saraswat.	2023	This paper proposes a novel approach for effective anomaly detection in surveillance footage using Attention Residual Long Short-Term Memory (ARLSTM) networks. The proposed ARLSTM architecture integrates attention mechanisms, residual connections, and LSTMs to improve the detection accuracy of anomalies in surveillance footage.	ARLSTM networks, especially when integrating attention mechanisms and residual connections, is computationally intensive. This results in longer training times and increased resource requirements during inference.	2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)
10	A Review On Anomalous Movement Fraudulent Detection in Exam Hall Priyanka Padhiyar, Rashmin Prajapati	2022	This paper focus on how the Abnormal behaviour detection system help in discovery of malicious activity and assist in the decrease of such actions.	However, challenges related to scalability, computational requirements, and potential biases in the dataset used may be considered disadvantages.	2nd International Conference on Artificial Intelligence and Smart Energy (ICAIS)

ADVANTAGES OF CURRENTLY EXISTING SYSTEMS

- **Pre-trained Models:** Pre-trained models are available for many existing systems, allowing users to leverage transfer learning and quickly adapt these models to new tasks with minimal additional training.
- **Comprehensive Datasets:** Many existing systems have been tested on large datasets, making their models reliable and able to handle various situations effectively.
- **Automation :** Enables the automation of surveillance processes, reducing the reliance on manual monitoring.

DISADVANTAGES OF CURRENTLY EXISTING SYSTEMS

- **False Positives :** One of the major challenges is the potential for false positives, where normal behaviour is mistakenly identified as abnormal, leading to unnecessary alerts.
- **Cost of Implementation :** Implementation of sophisticated systems can be expensive, involving the cost of hardware, software, and ongoing maintenance.
- **Technical Limitations :** Technical limitations, such as environmental conditions, lighting, and camera angles, may affect the accuracy of abnormal activity detection.

PROBLEM STATEMENT

- In modern surveillance systems, the sheer volume of video data generated poses a significant challenge for effective monitoring and threat detection.
- In densely populated and crowded public spaces traditional video surveillance systems face heightened challenges in detecting anomalies amidst the complexity of normal activities.
- The problem at hand is to develop an advanced video anomaly detection system specifically tailored for crowded environments.
- The objective is to identify unusual events or behaviors within dense crowds that may indicate security threats, safety concerns, or abnormal occurrences.

PROPOSED SOLUTION

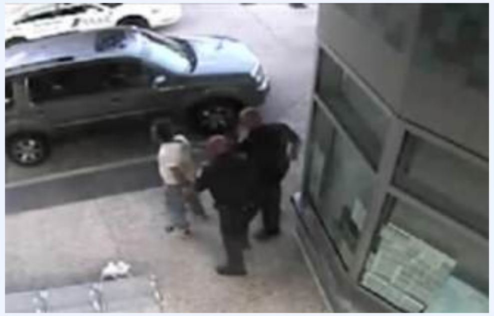
- We use Convolutional Neural Network (CNN) for abnormal activity detection as it provides a powerful and flexible framework leveraging their ability to automatically learn hierarchical features and capture spatial-temporal patterns.
- The steps taken includes :
  1. Data Collection : Collect video data from the UCF Crime Dataset which contains 14 classes.
  2. Aggressive Motion Detection : Apply Lucas-Kanade optical flow techniques to identify aggressive motions taking place in a video.
  3. Apply Grayscale Transformation : Convert the aggressive frames to grayscale, reducing the complexity of subsequent processing steps.

4. Apply Canny Edge Detection : Enhance features by applying Canny edge detection to the grayscale frames, highlighting contours and edges.
5. Frame Storage : Save the preprocessed frames in a designated folder for later use in training and testing.
6. Data Splitting : Split the saved frames into training and testing sets.
7. Model Training : Train a CNN model on the training set.
8. Abnormal Activity Detection : Apply the trained CNN model to new videos, detecting abnormal activities in each frame.
9. Outcome Evaluation : Evaluate the model's performance using metrics such as accuracy, confusion matrix, ROC curve and AUC score.

DATASET

- The dataset used is UCF Crime Dataset.
- It has a total of 13 classes of abnormal activities and 1 class of normal activities.


Abuse


Arrest


Arson


Assault


Burglary


Explosion


Fighting


Road Accidents

Robbery

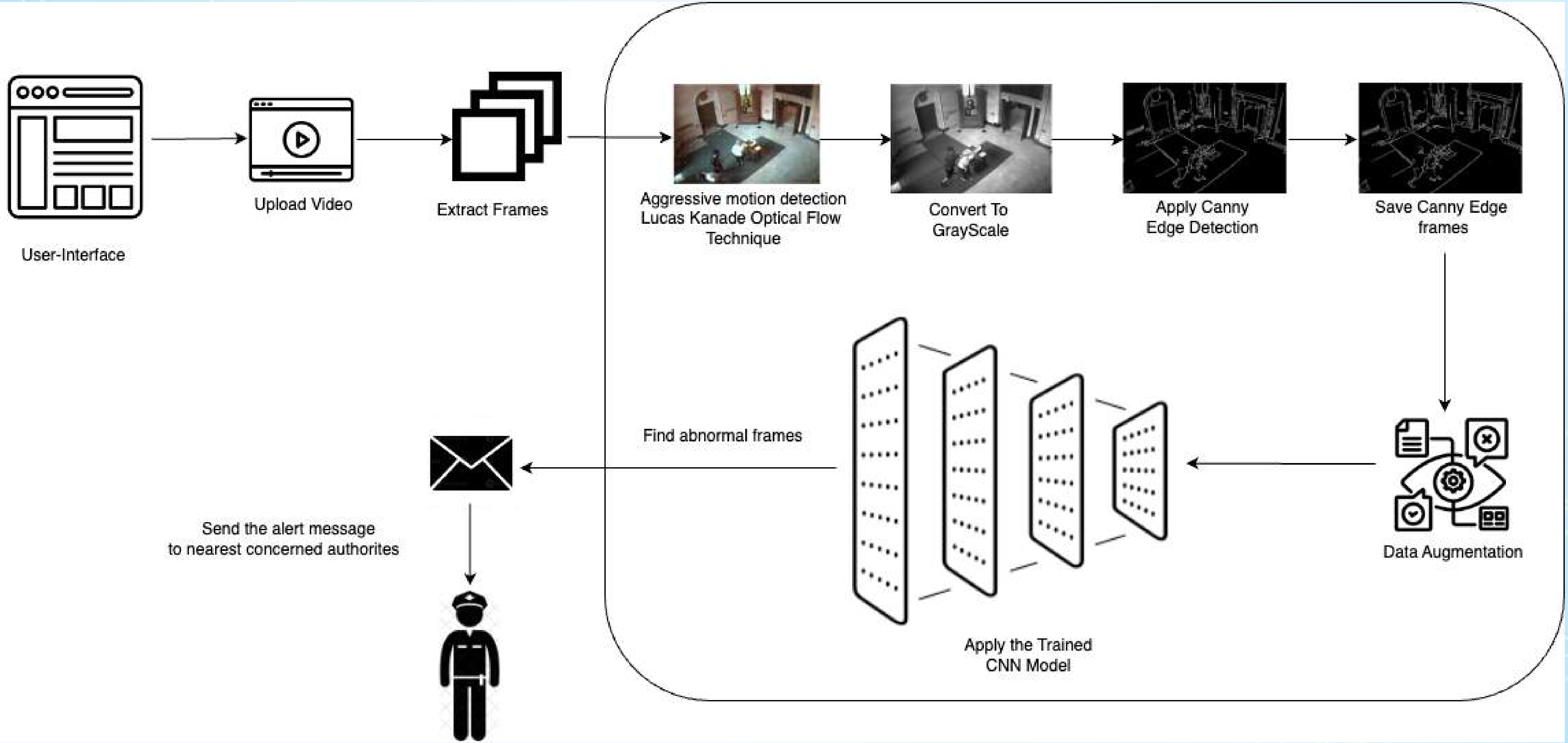
Shooting

Shop Lifting

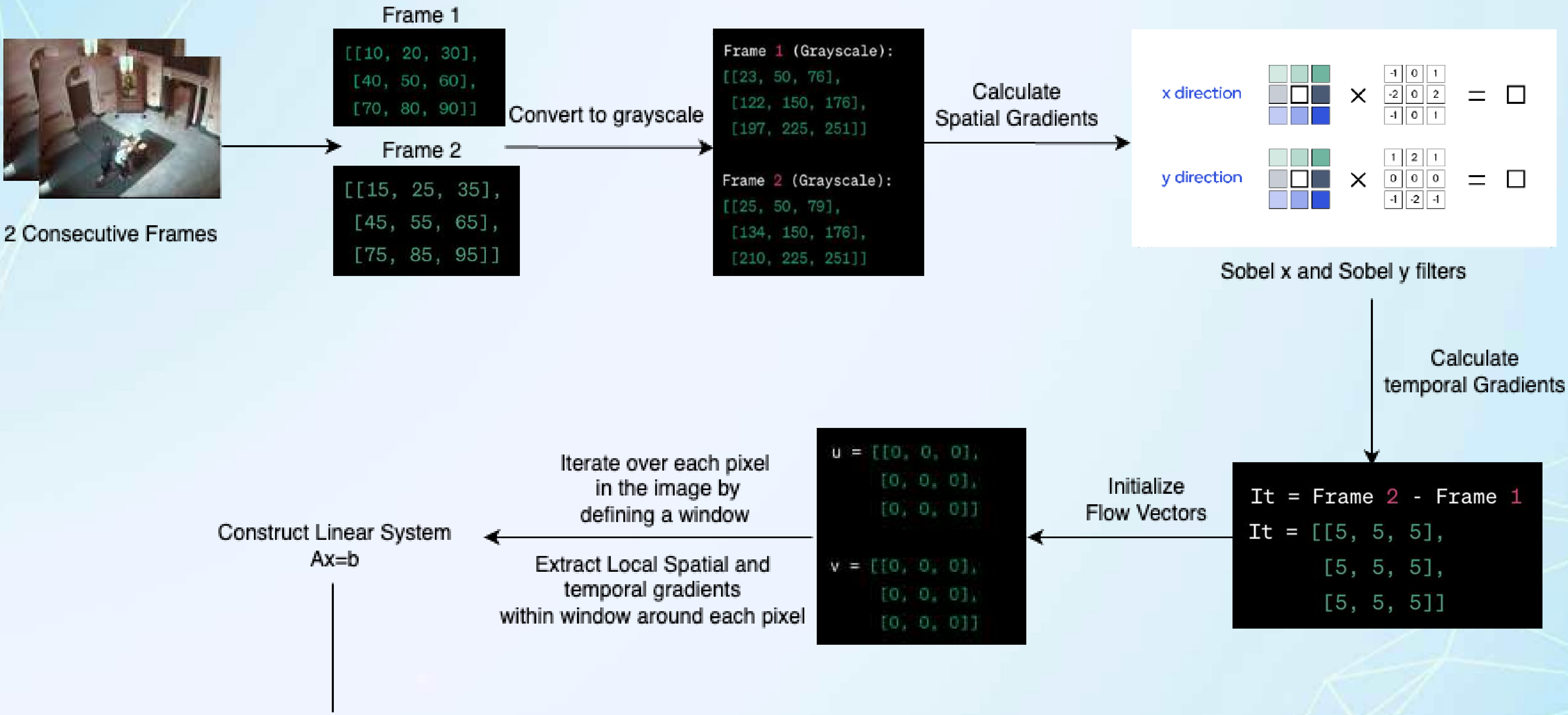
Stealing

Vandalism

HIGH LEVEL DESIGN



LUCAS – KANADE ALGORITHM



```
Ix = [[-10, 0, 10],
      [-20, 0, 20],
      [-10, 0, 10]]

Iy = [[-10, -20, -10],
      [0, 0, 0],
      [10, 20, 10]]

It = [[5, 5, 5],
      [5, 5, 5],
      [5, 5, 5]]
```

3x3 window with Ix,Iy and It

and

```
Ix_window = [-10, 0, 10, -20, 0, 20, -10, 0, 10]
Iy_window = [-10, -20, -10, 0, 0, 0, 10, 20, 10]
It_window = [5, 5, 5, 5, 5, 5, 5, 5, 5]
```

Gradients of specific pixel in the window

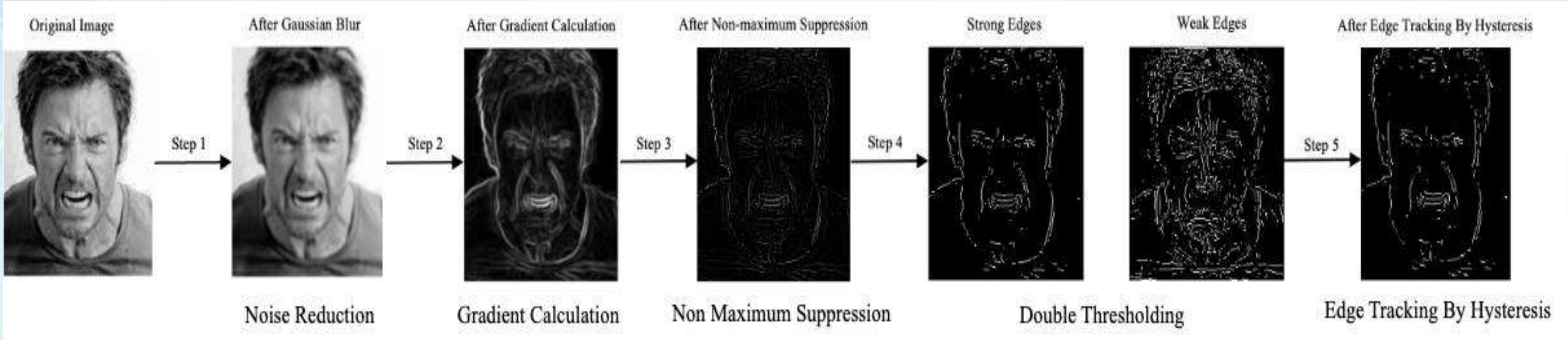
```
A = [[-10, -10],
     [0, -20],
     [10, -10],
     [-20, 0],
     [0, 0],
     [20, 0],
     [-10, 10],
     [0, 20],
     [10, 10]]

b = [[-5],
     [-5],
     [-5],
     [-5],
     [-5],
     [-5],
     [-5],
     [-5],
     [-5]]
```

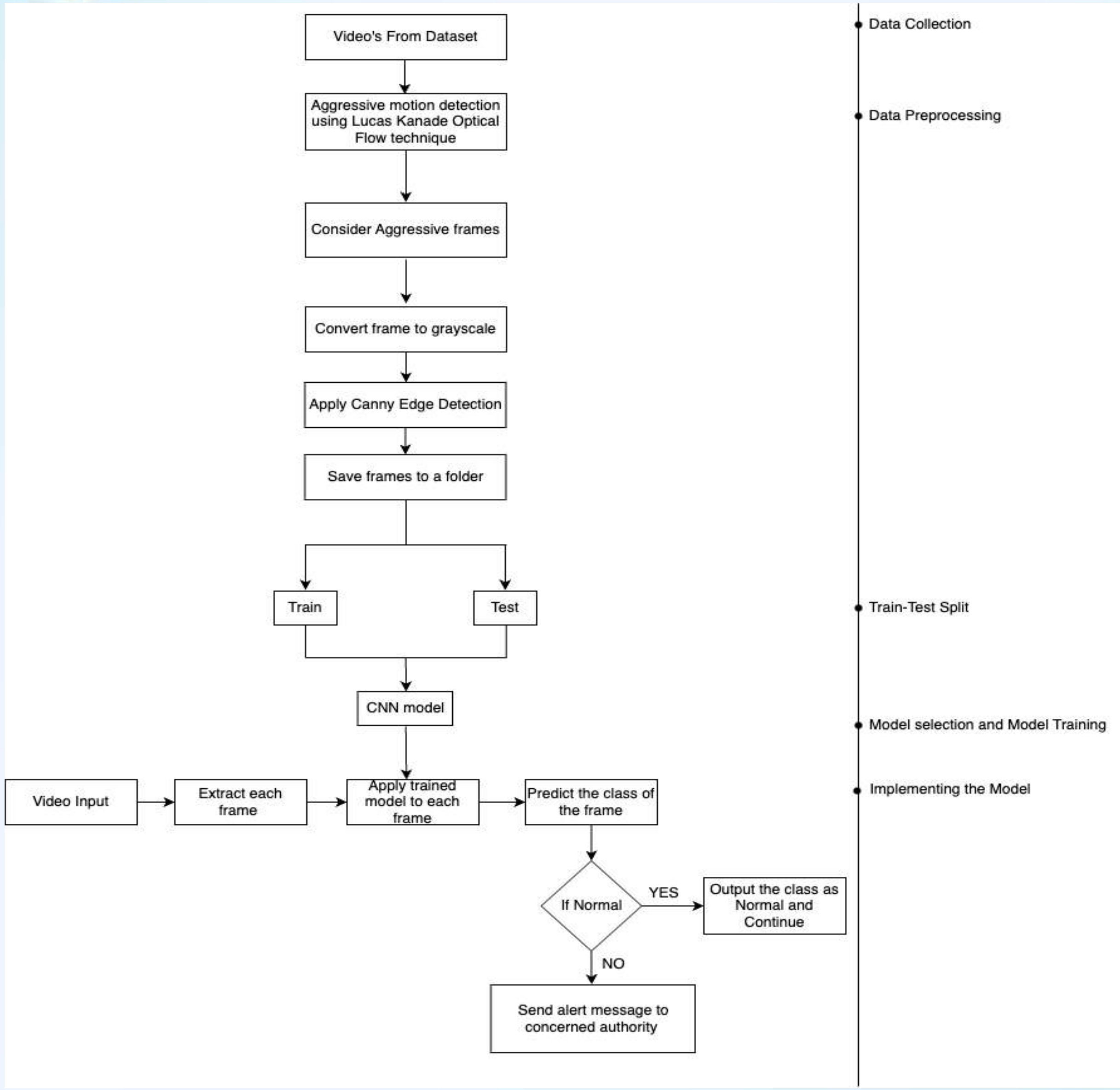
Ax=b

1. Solve the Linear System to find flow vector uv and update flow vectors u and v
2. Find the magnitude and angle from vectors u and v

Canny Edge Detection



SYSTEM ARCHITECTURE



IMPLEMENTATION

DATA PREPROCESSING :

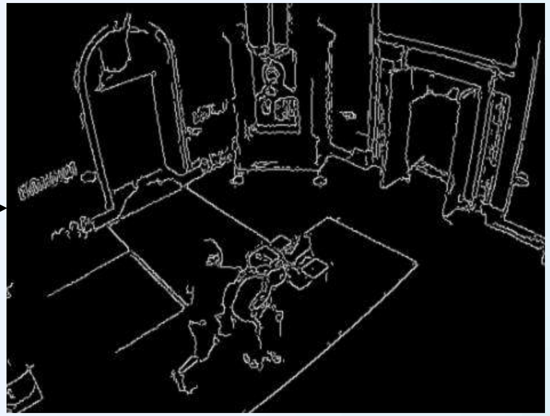
Step 1: Detect Aggressive motions from videos using Lucas-Kanade Optical Flow algorithm

Video

Aggressive Frames

Step 2: Apply Canny edge detection to Aggressive Frames

Aggressive Frames



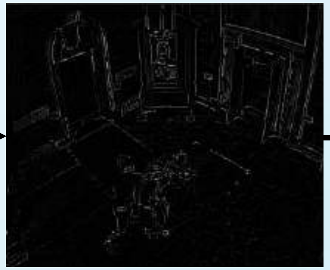
After applying Canny Edge Detection
On Aggressive Frames

CANNY EDGE DETECTION PROCESS:

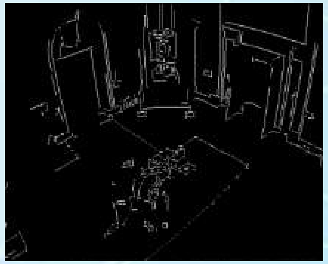
Gray Scale Conversion

Gaussian Blur

Gradient Calculation



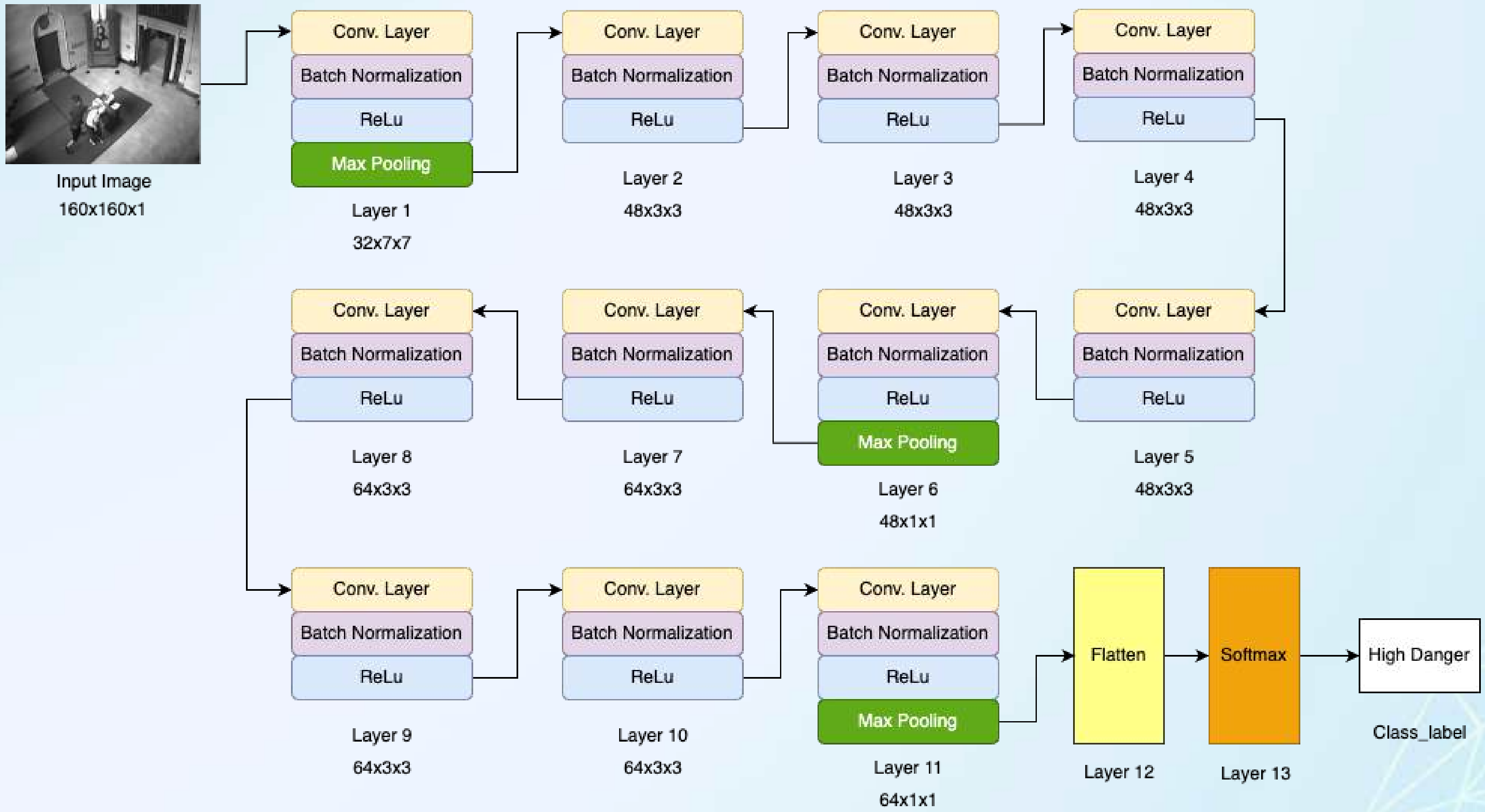
Non-Maximum Suppression



Final edges by hysteresis

Step 3: Save Canny edge applied files to a folder and split folder into test and train.

DEFINE CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE:



- The proposed model contains a total of 13 layers excluding input layer.
- In each layer convolution layer is followed by batch normalization and ReLu Layer.
- Along with those Max Pooling layer is also present in layers 1,6 and 11.
- Then Flatten is used to convert the output of layer 11 to a 1-D large vector or array.
- Its output is then passed to Softmax layer which classifies the input image into one of the defined classes.

DEFINE IMAGE DATA GENERATORS:

- These generators are created for train data that preprocess images and apply various data augmentation techniques such as rescaling, rotation, flipping, shifting, etc on-the-fly.

LOADING TRAINING AND TESTING DATA:

- Training and testing data are loaded using the flow_from_directory method of ImageDataGenerator.

MODEL COMPILATION :

- The model is compiled with categorical cross-entropy loss and the Adam optimizer. The accuracy metric is specified to monitor during training.

MODEL TRAINING :

- The model is trained using the **fit** method. The training data and validation data are provided, along with the number of epochs and steps per epoch.

MODEL EVALUATION :

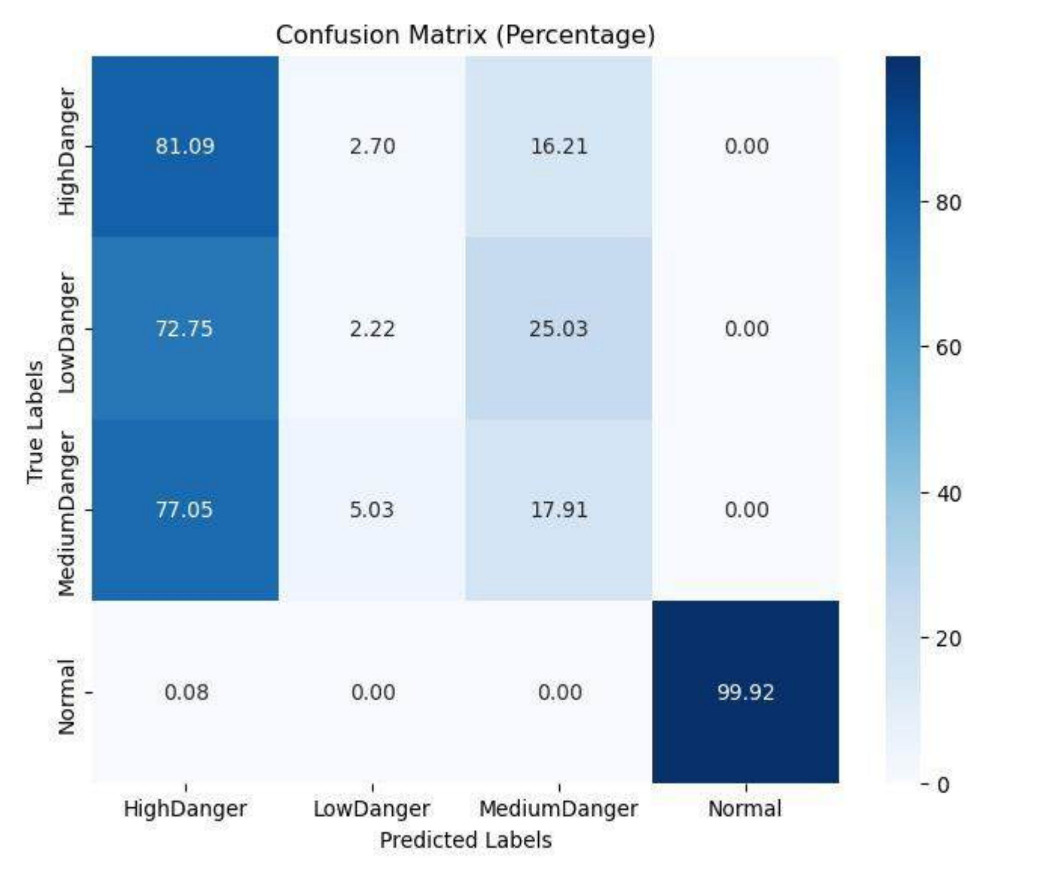
- The trained model is evaluated on the testing data using the **evaluate** method.

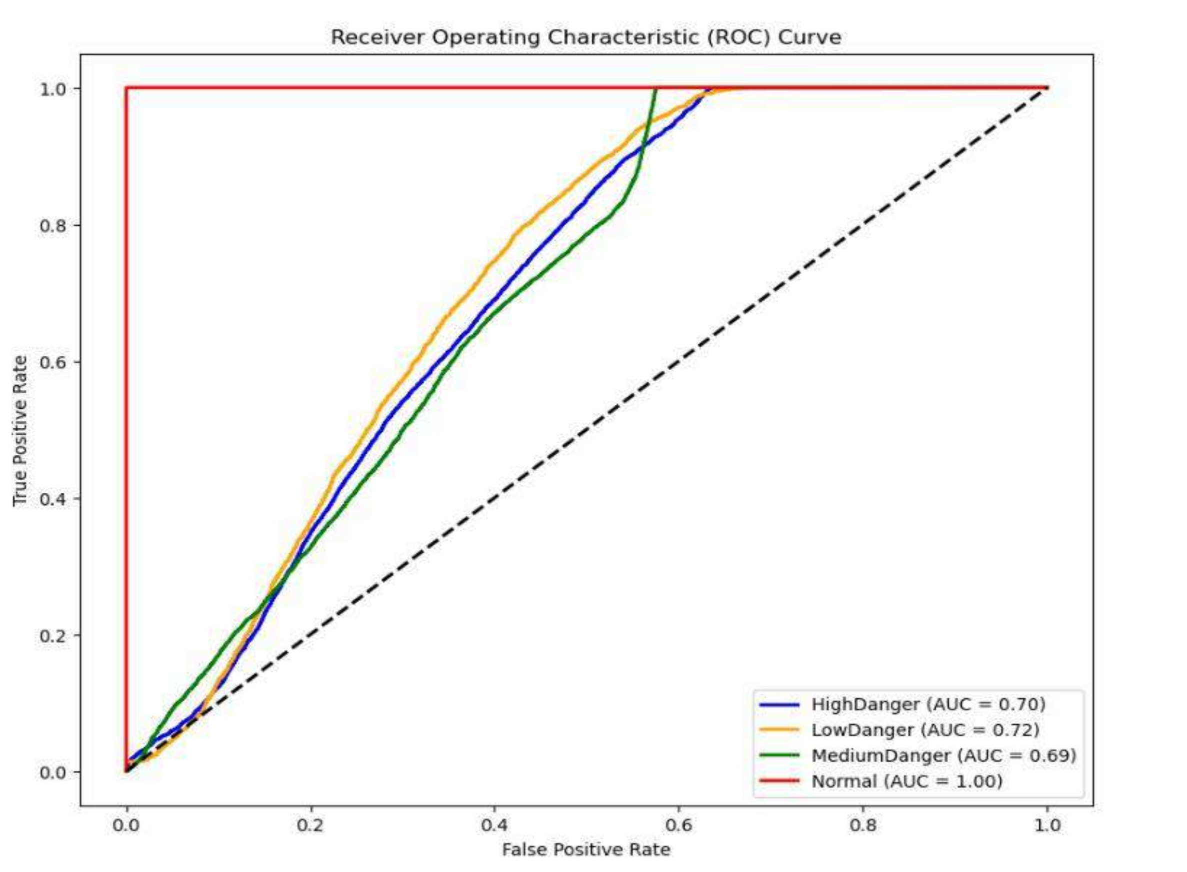
SAVE THE MODEL :

- The trained model is saved using the **save** method.

PERFORMANCE METRICS :

- **ROC Curve:** The Receiver Operating Characteristic (ROC) curve illustrates a classifier's performance by plotting the true positive rate against the false positive rate at various threshold settings. The Area Under the ROC Curve (AUC-ROC) serves as a measure of the model's ability to distinguish between classes.
- **Accuracy of CNN Model:** Accuracy measures the proportion of correctly classified instances out of the total instances. It's a simple metric that indicates how often the model's predictions match the actual labels.





- AUC score obtained – 84.54%
- Accuracy of the model – 72.03%

Classification Report

Classification Report:

	precision	recall	f1-score	support
HighDanger	0.35	0.81	0.49	5225
LowDanger	0.11	0.02	0.04	2793
MediumDanger	0.46	0.18	0.26	7413
Normal	1.00	1.00	1.00	5937

Model Accuracy

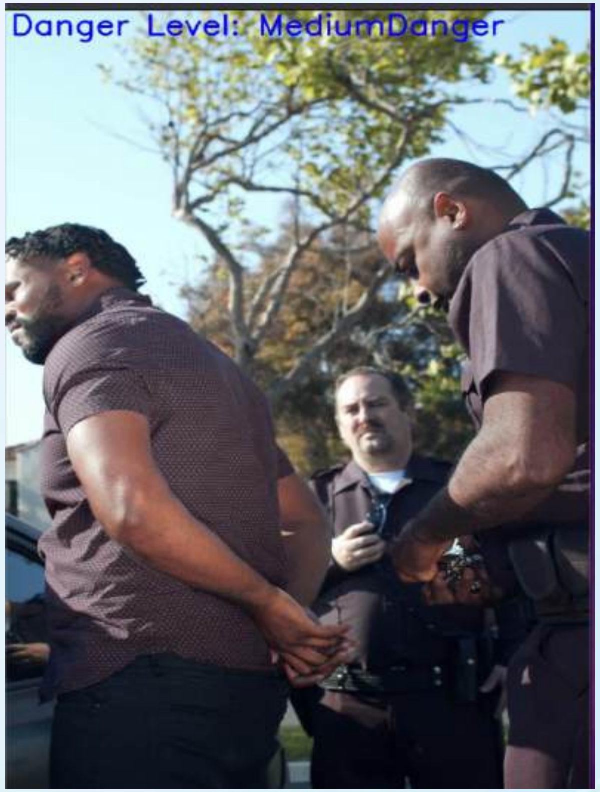
```
Found 21046 images belonging to 4 classes.
658/658 [==============================] - 458s 691ms/step - loss: 0.6375 - accuracy: 0.7203
Model Accuracy: 72.03%
```

APPLY THE SAVED MODEL TO THE NEW VIDEO :

- Define the class mapping using dictionary in python – high_danger : 3, medium_danger : 2, low_danger : 1, normal : 0.
- Extract each frame from the video and for each frame:
    - Resize the frame to 160x160 px.
    - Convert it to grayscale.
    - Convert grayscale frame to numpy array.
    - Expand dimension of frame so that it is compatible with model.
    - Make predictions using the saved model.
    - Display the predicted value or the key on the screen.
- If predicted value is 2 or 3 continuously for few seconds then send alert message to concerned authority.(Using twilio, a cloud communication platform).

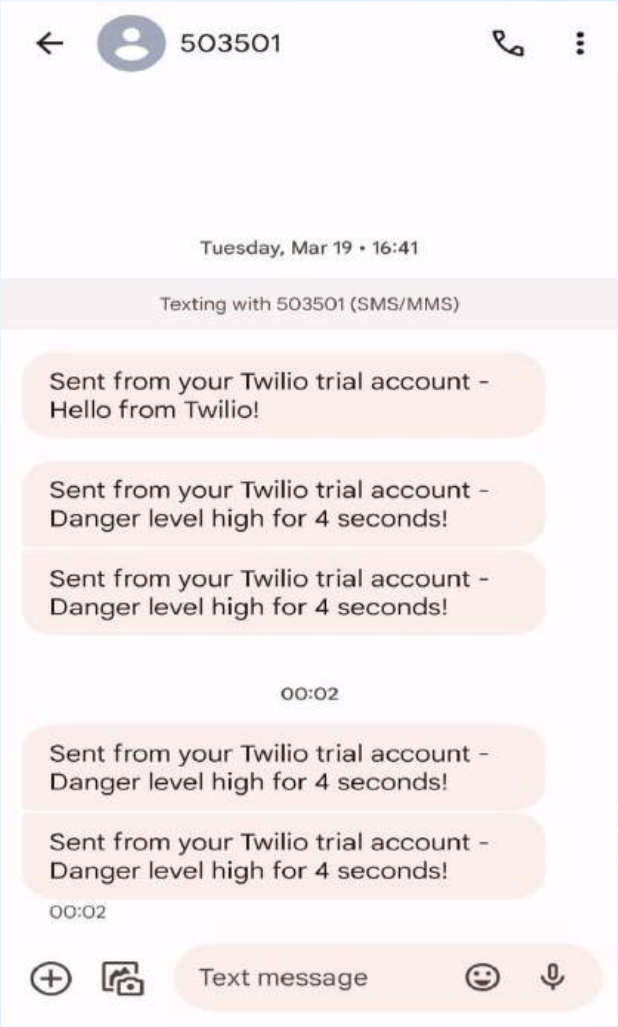
RESULTS:

High_Danger



Medium_Danger

Normal



Alert message from Twilio

APPLICATIONS:

- **Security and Surveillance**: Detecting suspicious or aggressive activities in security cameras.
- **Public Safety**: Monitoring public spaces for signs of violence or abnormal behavior.
- **Healthcare**: Identifying falls or other dangerous events in healthcare settings.
- **Smart Cities**: Enhancing urban safety through automated monitoring.
- **Retail and Commerce**: Identifying theft or other suspicious activities in stores.

FUTURE ENHANCEMENTS:

- **Resampling**: Resampling methods can help balance the class distribution, leading to improved model generalization.
- **Advanced CNN Architectures:** Experimenting with more complex CNN architectures, including deeper networks or different layer configurations, could improve the model's ability to capture intricate patterns and reduce misclassification rates.
- **Transfer Learning:** Leveraging pre-trained models through transfer learning can offer improved feature extraction and reduce training time, leading to better performance.

REFERENCES

[1] Zeng, Xianlin, et al. "A hierarchical spatio-temporal graph convolutional neural network for anomaly detection in videos." IEEE Transactions on Circuits and Systems for Video Technology 33.1 (2021): 200-212.

[2] D. -G. Lee, H. -I. Suk, S. -K. Park and S. -W. 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.

[3] Direkoglu, Cem. "Abnormal crowd behavior detection using motion information images and convolutional neural networks." IEEE Access 8 (2020): 80408-80416.

[4] R. Sittalatchoumy, S. E. P. Pushpa, R. Karthick, M. A. Rafiq Azharudheen and K. V. Dhiyanesh, "Transfer Learning based Video Surveillance for Abnormal Activity Detection," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023.

[5] S. Nikkath Bushra, G. Shobana, K. Uma Maheswari and N. Subramanian, "Smart Video Survillance Based Weapon Identification Using Yolov5," 2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC), Chennai, India, 2022.

[6] Bajgoti, Arpit, et al. "SwinAnomaly: Real-Time Video Anomaly Detection using Video Swin Transformer and SORT." IEEE Access (2023).

[7] R. Hasib, A. Jan and G. M. Khan, "Real-Time Anomaly Detection for Smart and Safe City Using Spatiotemporal Deep Learning," 2022 2nd International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, 2022.

[8] Dileep, Arjun S., et al. "Suspicious human activity recognition using 2d pose estimation and convolutional neural network." 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). IEEE, 2022.

[9] J. Suganthi and V. Abineshkumar, "Revolutionizing Anomaly Detection in Surveillance Footage with ARLSTM," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023.

[10] P. Padhiyar and R. Prajapati, "A Review On Anomalous Movement Fraudulent Detection in Exam Hall," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022.

Thank You