Department of Computer Science and Engineering
Vimal Jyothi Engineering College
Chemperi

A Real-time Violence Recognition System for Surveillance Cameras

MEMBERS:

Angel John(VML20CS038)
Ann Maria George(VML20CS044)
Jithina Raj P(VML20CS091)
Vismaya Hemanth Nambiar
(VML20CS181)

GUIDE:

Mr. Abhiram P

CONTENTS

AREA OF SELECTION

DEEP LEARNING

Deep learning is a subset of artificial intelligence that employs artificial neural networks to automatically learn from extensive datasets. These neural networks, with their intricate layers of interconnected nodes, are particularly effective at handling complex tasks such as image recognition, speech analysis, and natural language understanding.

TOPIC:

A Real-time Violence Recognition System for Surveillance Cameras using CNNs

ABSTRACT

- The growing demand for improved security and safety in public spaces, commercial establishments, and critical infrastructures has spurred interest in the development of efficient violence detection systems for automated surveillance applications.
- Project focuses on the design and implementation of a cutting-edge violence detection solution that harnesses the power of deep learning and computer vision technologies.
- Project advances automated surveillance tools, providing a promising solution for addressing security challenges in our dynamic world, with significant potential for applications in law enforcement, public safety, and critical infrastructure protection.

INTRODUCTION

- In today's increasingly complex and security-conscious world, the need for advanced surveillance technologies that can proactively identify and respond to threats has become paramount.
- This research project delves into the development of an efficient violence detection system specifically tailored for automated surveillance applications.
- By harnessing the power of cutting-edge technologies such as deep learning and computer vision, we aim to provide a reliable and rapid response mechanism to enhance security measures and contribute to safer environments.

PROBLEM DEFINITION

- The abundance of video data from surveillance cameras, combined with the limitations of human real-time analysis, has created a delay in detecting events.
- The project explores the use of advanced networks and pre-trained models to automatically identify violent actions in surveillance footage.

SCOPE OF THE SYSTEM

- To enhance the precision and speed of identifying violent incidents within surveillance footage, particularly in the context of managing large volumes of video data and addressing compression challenges during remote processing.

OBJECTIVE

- Develop an efficient violence detection system for surveillance videos.
- Improve accuracy and speed of identifying violent actions, even when faced with compression artifacts in remote server processing.
- A critical component of the system is the ability to alert relevant authorities when violent actions are detected.

LITERATURE REVIEW

The papers selected for the literature survey are:

- Toward Fast and Accurate Violence Detection for Automated Video Surveillance Applications
- Low-Cost CNN for Automatic Violence Recognition on Embedded System
- Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems
- Efficient Anomaly Detection in Crowd Videos Using Pre-Trained 2D Convolutional Neural Networks

Paper 1: Toward Fast and Accurate Violence Detection for Automated Video Surveillance Applications

BRIEF SUMMARY

- The study uses seven different datasets commonly used in violence detection research to facilitate comparisons with other methods.
- These datasets include videos of violence in crowds, hockey fights, scenes from action movies, real-life street fights, abuse and so on .
- The researchers finetune a model called X3D-M for violence detection.

Paper 1: Toward Fast and Accurate Violence Detection for Automated Video Surveillance Applications

- A fully functional stand-alone system that implements the proposed methods for automated violence detection is presented.
- The database combines and extends seven existing video databases: Crowd Violence,Hockey Fights,Movie Fights,Real Life Violence Situations,Real-World Fight-2000,UCF-Crime Selected,XD-Violence Selected .

Paper 2: Low-Cost CNN for Automatic Violence Recognition on Embedded System

BRIEF SUMMARY

- This paper evaluates how mobile CNNs can perform the task of automatic violence recognition.
- The work of violence recognition is typically a binary classification problem, that is, it has only two classes: violence and nonviolence.

Paper 2: Low-Cost CNN for Automatic Violence Recognition on Embedded System

- A "third class" was created in the sigmoid function, that is, instead of using only a threshold, a zone of uncertainty was delimited.
- The study uses and compares different models of Mobile Convolutional Neural Networks, to avoid overfitting and improve the model, generalization, and traditional data augmentation techniques were applied.

Paper 3: Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems

BRIEF SUMMARY

- A novel approach called Tuna Swarm Optimization with Deep Learning Enabled Violence Detection (TSODL-VD) for real-time violence detection in surveillance videos is introduced.
- This technique utilizes a residual-DenseNet model to extract features from video frames and employs a stacked autoencoder (SAE) classifier to distinguish between violent and non-violent events.
- It also incorporates the TSO protocol as a hyperparameter optimizer for the model. Experimental results on a benchmark violence dataset show that TSODL-VD outperforms existing methods, providing precise and fast violence detection.

Paper 4:Efficient Anomaly Detection in Crowd Videos Using Pre-Trained 2D Convolutional Neural Networks

BRIEF SUMMARY

- The continuous tracking of surveillance videos is not workable for humans. Therefore, it is focused on proposing efficient methods for autonomous monitoring systems that observe abnormal crowd activities.
- This approach eliminate the need for training 3D CNNs using a video dataset. In particular, it suggests to exploiting 2D CNNs pre-trained on images for learning both spatial and temporal information.
- This approach enables the use of many readily available image-trained 2D CNNs,and also significantly reduces the requirements for computational and memory resources.

Paper 4:Efficient Anomaly Detection in Crowd Videos Using Pre-Trained 2D Convolutional Neural Networks

- The former stream of the 2 stream CNN learns the spatial characteristics of objects in the scene, whereas the latter trains on indispensable temporal features.
- In each frame,this method discards the patches that contain little motion information. So, it uses data from only selected volumes of interest (VOIs) carrying rich information pertaining to motion. Therefore, regions where no pedestrians appear are discarded.
- Second, to extract and represent the temporal information, a low computational cost method stacked grayscale 3-channel image (SG3I) is applied instead of optical flow.

CONSOLIDATED TABLE

PAPER	PROBLEM STATEMENT	TECHNOLOGIES	FINDINGS
Toward Fast And Accurate Violence Detection for Automated Video Surveillance Applications	Automated violence detection in surveillance footage in real time	3D convolution,Human Activity Recognition (HAR),CNN,X3D-M	Computationally light, adaptable to new scenarios, good classification accuracy
Low-Cost CNN for Automatic Violence Recognition on Embedded System	Low cost CNN to automatically recognize suspicious events to alert and assist monitoring process with reduced deployment cost	LSTM,CNN, SQUEEZENET	Mobile CNN models achieve the possible result with limited processing power
Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems	Tuna Swarm Optimization with Deep Learning Enabled Violence Detection(TSODL-VD) technique to classify violent actions in surveillance videos	TSODL-VD,stacked autoencoder (SAE) classifier, residual-DenseNet model	Precise and rapid detection outcomes
Efficient Anomaly Detection in Crowd Videos Using Pre-Trained 2D Convolutional Neural Networks	Detection of crowd abnormal behaviors by using spatial and temporal information got from videos	2D CNN, stacked grayscale 3-channel image (SG3I)	reduces the requirements for computational and memory resources

REQUIREMENT SPECIFICATION

FUNCTIONAL REQUIREMENTS

- Scanning Capability
- Reporting
- Alerting
- Automation
- Integration
- Configuration
- Customization

NON-FUNCTIONAL REQUIREMENTS

- Real-time Processing
- Accuracy
- Scalability
- Reliability
- Security
- User Image
- Feedback Mechanism

SYSTEM HARDWARE REQUIREMENTS

- CPU: 2+ Cores, 2.46+ Ghz
- RAM: 6 GB or higher
- Disk: 25 GB + free space

ARCHITECTURE DIAGRAM

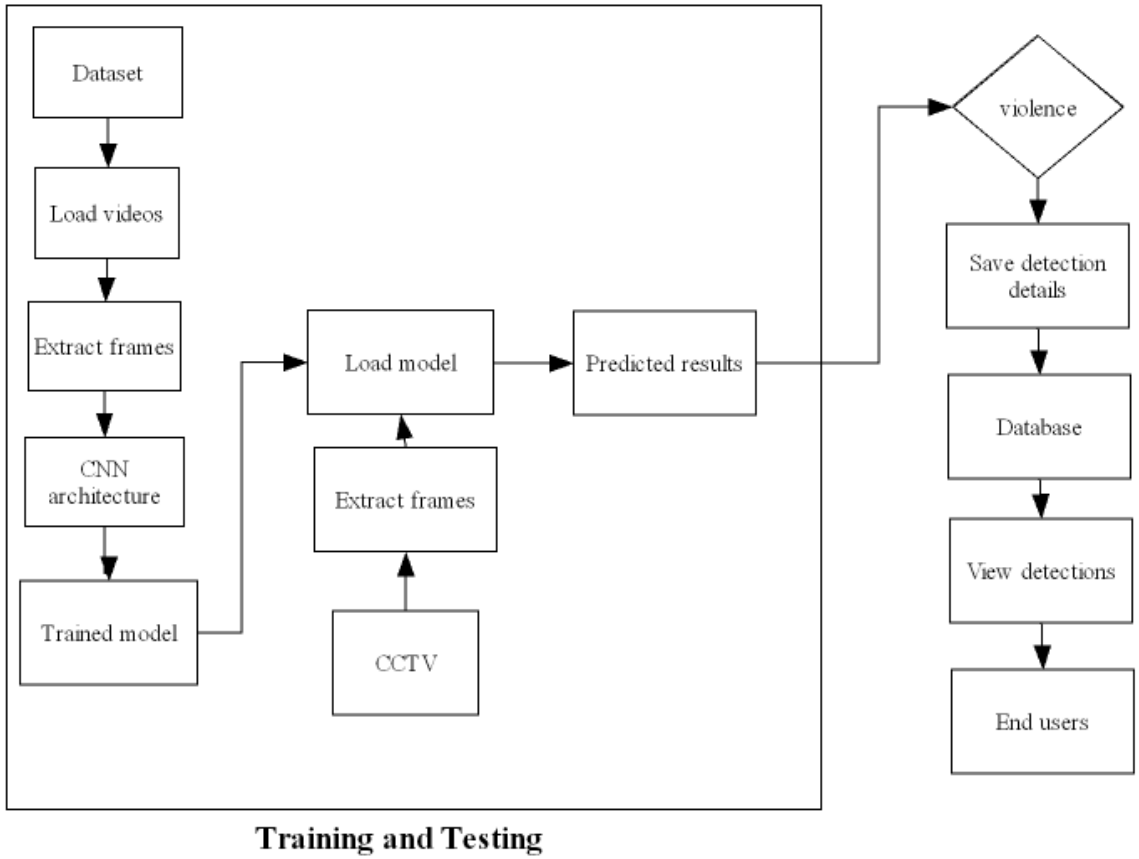
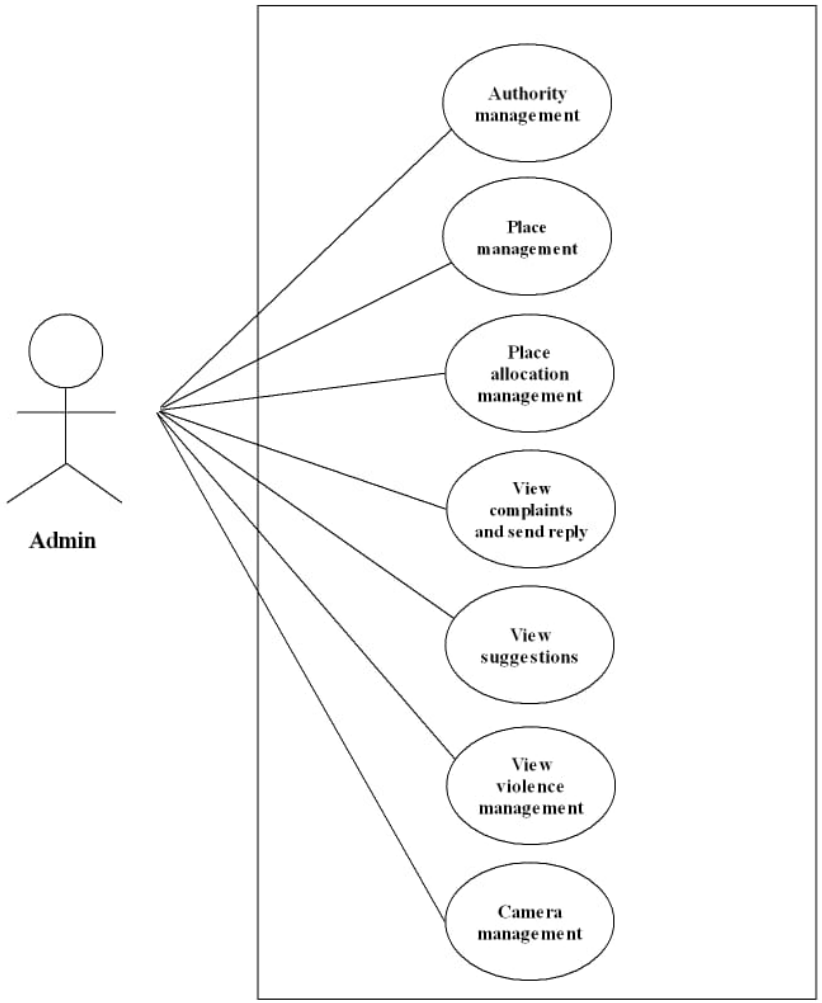


Figure: ARCHITECTURE DIAGRAM

USE CASE DIAGRAM



USE CASE DIAGRAM

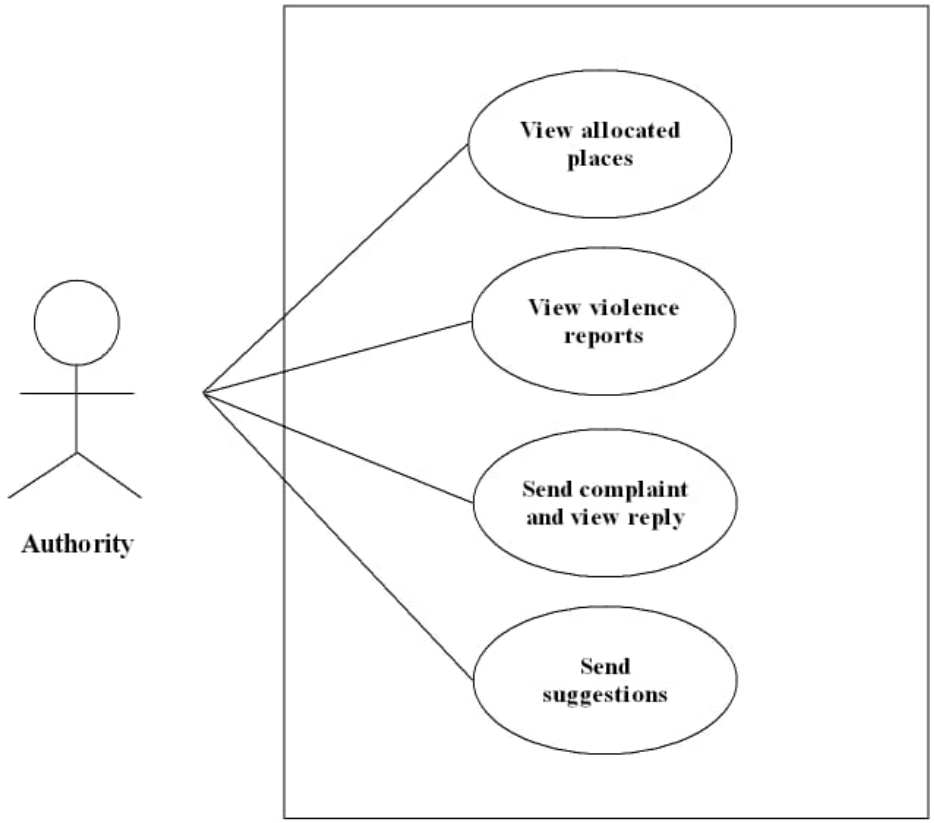


Figure: Use Case Diagram

USE CASE DIAGRAM

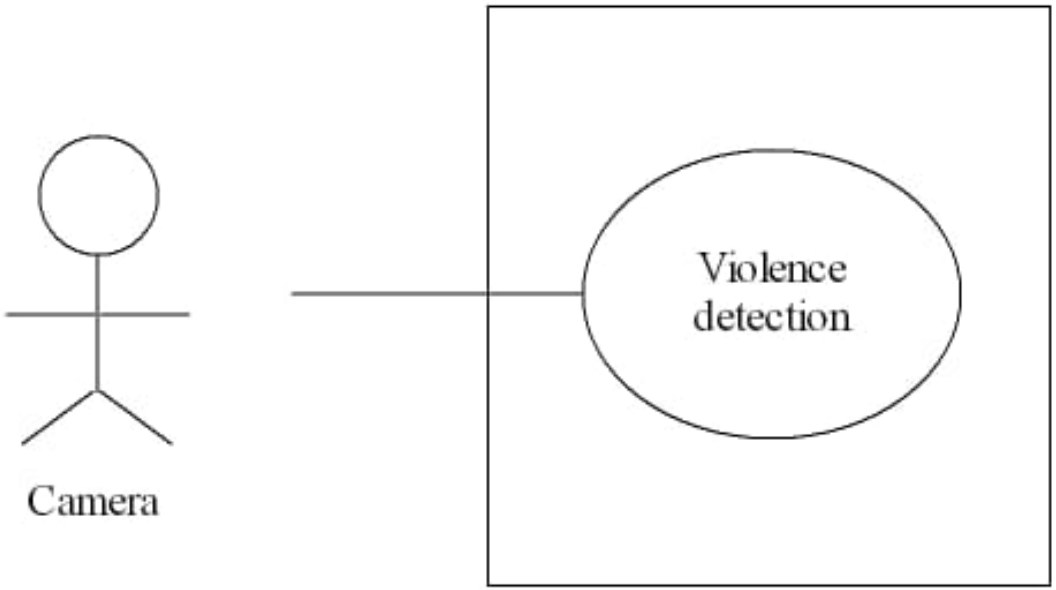


Figure: Use Case Diagram

DATA FLOW - LEVEL 0

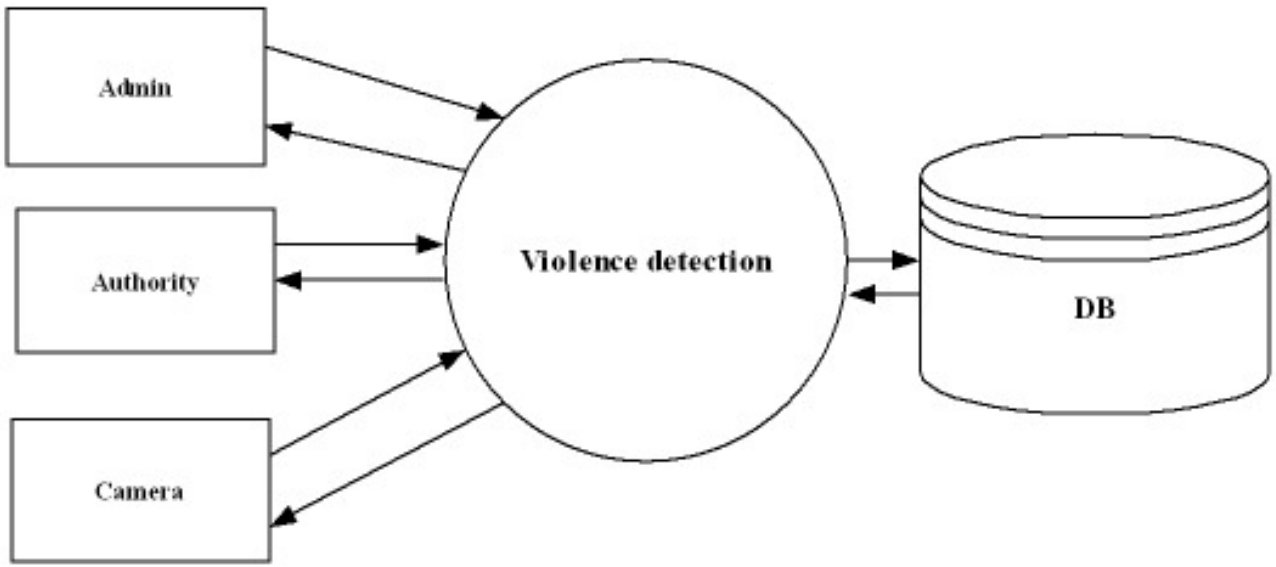
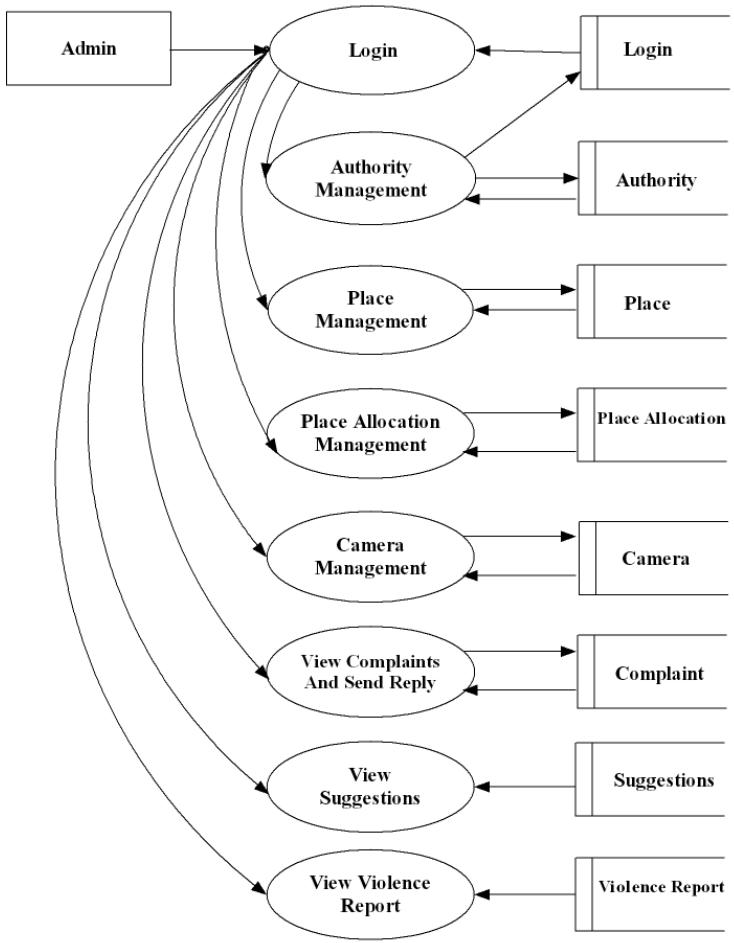


Figure: Data Flow Diagram

DATA FLOW - LEVEL 1.1



DATA FLOW - LEVEL 1.2

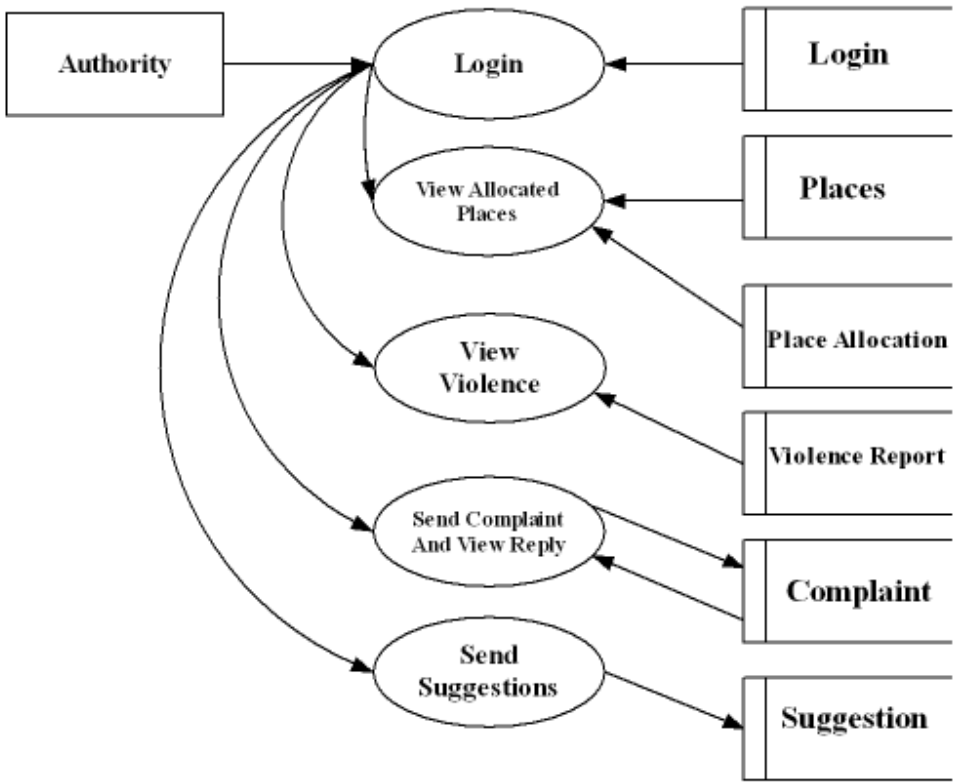


Figure: Data Flow Diagram

DATA FLOW - LEVEL 1.3

Figure: Data Flow Diagram

ER Diagram

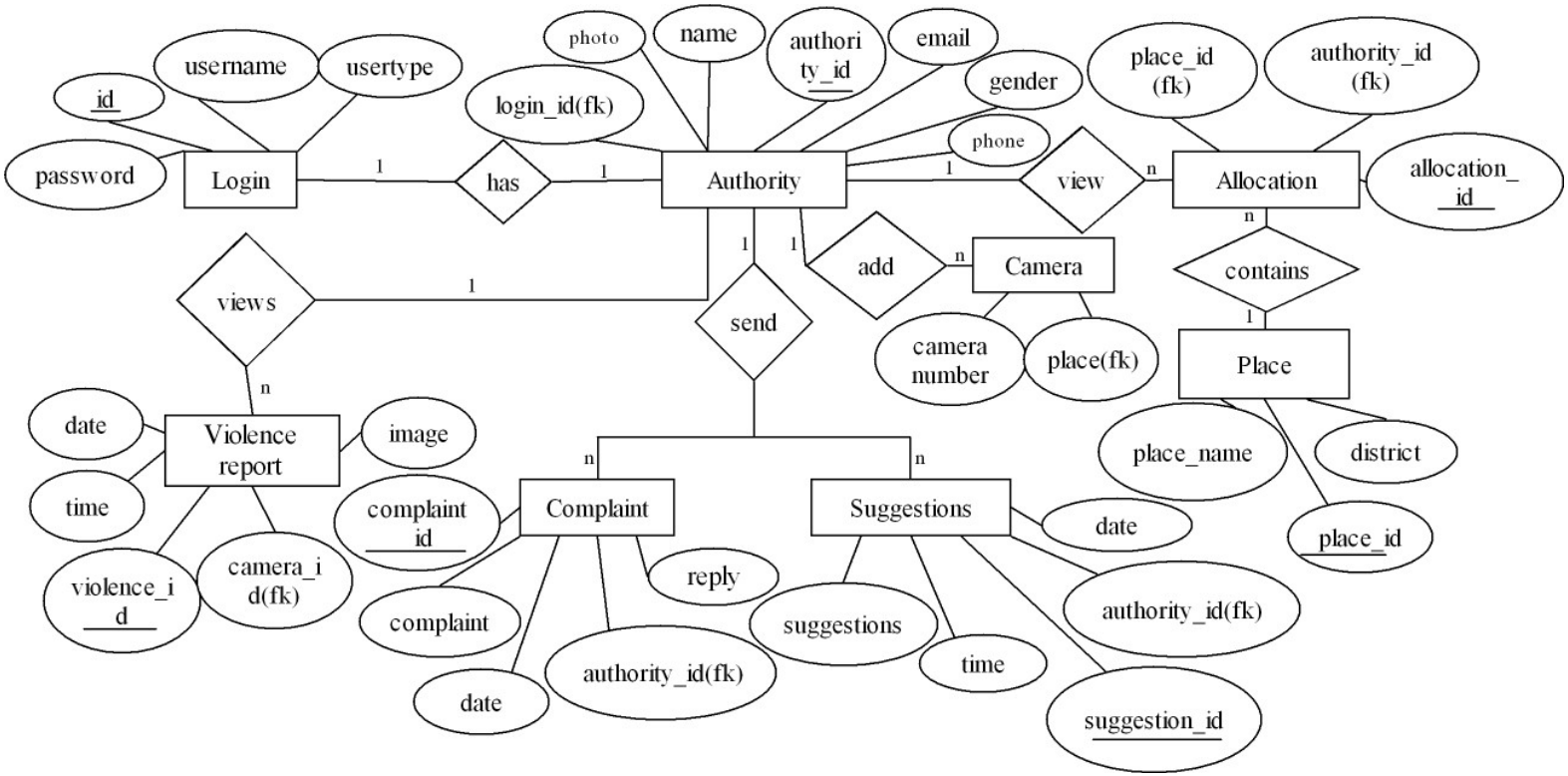


Figure: ER Diagram

FEASIBILITY STUDY

- **TECHNICAL FEASIBILITY:**
  The main technologies and tools that are associated with this project are: Pycharm,SQLyog, and Wamp Server. Most of these are open source and freely available and the technical skills required are manageable as there are well-described docs available. Hence this project is technically feasible.
- **ECONOMIC FEASIBILITY:**
  The system will have an associated hosting cost, which is affordable in this case. Also, since open source tools are used, the cost for development can be eliminated. Hence this project is economically feasible.

METHODS AND TECHNIQUES

- **Deep Learning :** A method in artificial intelligence (AI) that teaches computers to process data in a way that is inspired by the human brain.
- **CNN:** CNN stands for Convolutional Neural Networks (CNNs) are sophisticated algorithms designed for interpreting visual data, mimicking how the human brain perceives images for tasks like object recognition and classification.
- **Image Processing :** Image processing is the manipulation and analysis of digital images to enhance, extract information, or transform their appearance for various applications.
- **OpenCV :** OpenCV, a computer vision package that helps to read videos and also convert them into images is used.

TOOLS

- **PyCharm**
- **SQLyog**
- **WampServer**
- **Python Interpreter**

IMPLEMENTATION

- **Back-End designing :** The violence detection database contains tables needed for storing the data of admin, authority, and the camera such as allocation, complaints, login, places, suggestions, violence, etc.
  Training and testing completed.
- **Front-End designing:** The website model is completed.
- **Django project creation:**Creating basic structure and configuration files for a web application.
- **Routing:**Creating URLs to load and fetch HTML pages.

IMPLEMENTATION

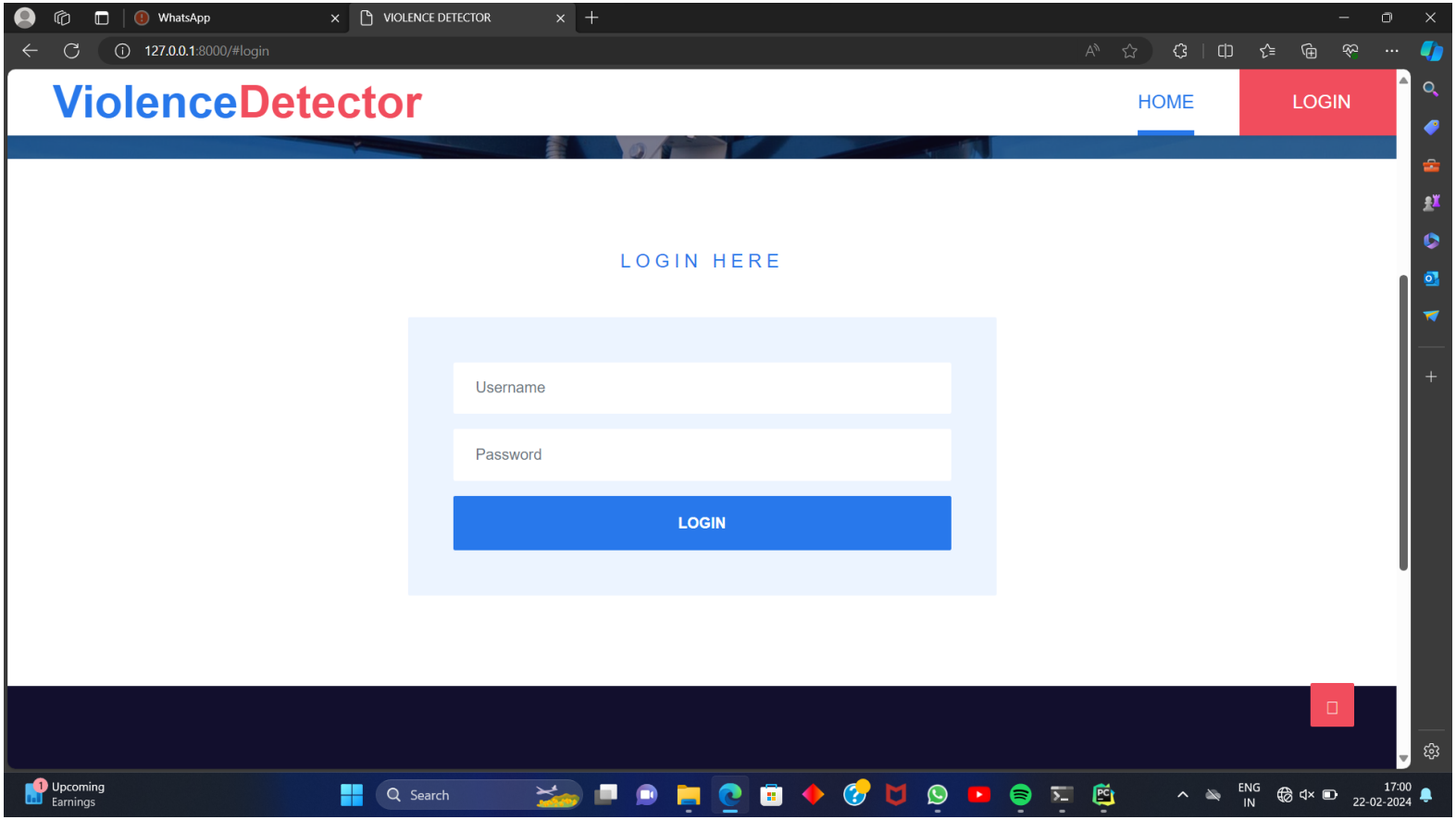


Figure: Login Page

IMPLEMENTATION

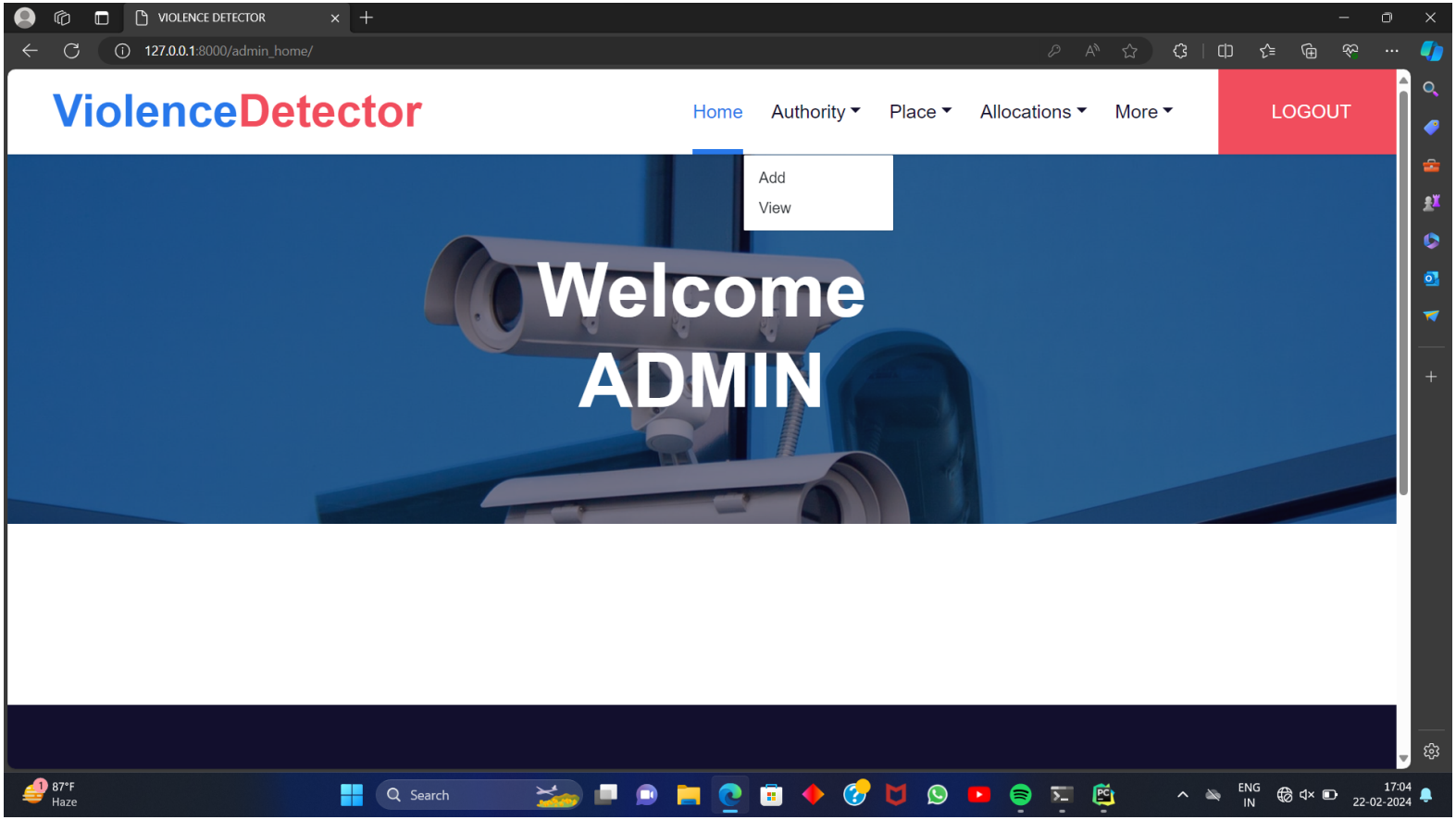


Figure: Admin Page

IMPLEMENTATION

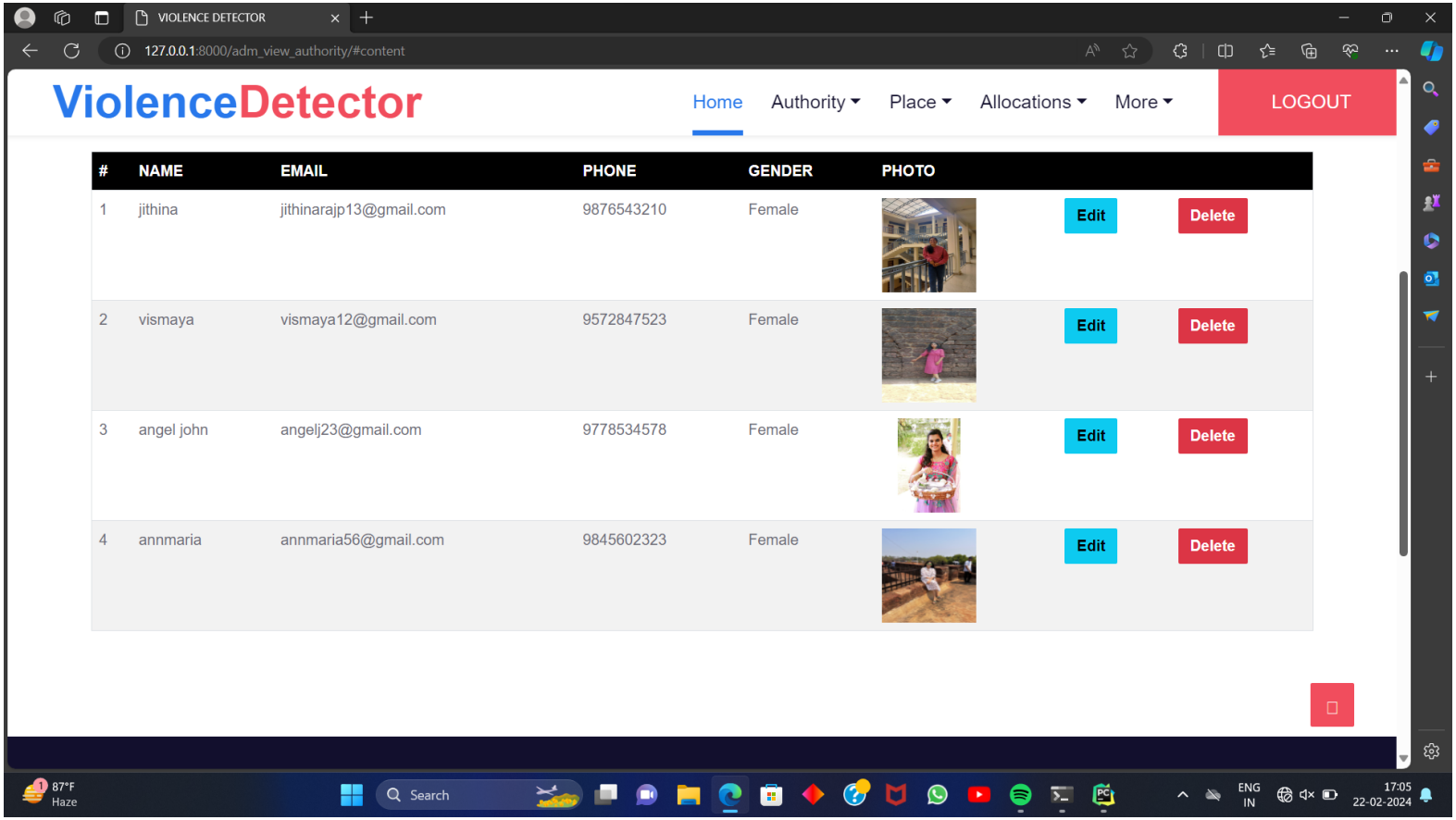


Figure: Added Authorities

IMPLEMENTATION

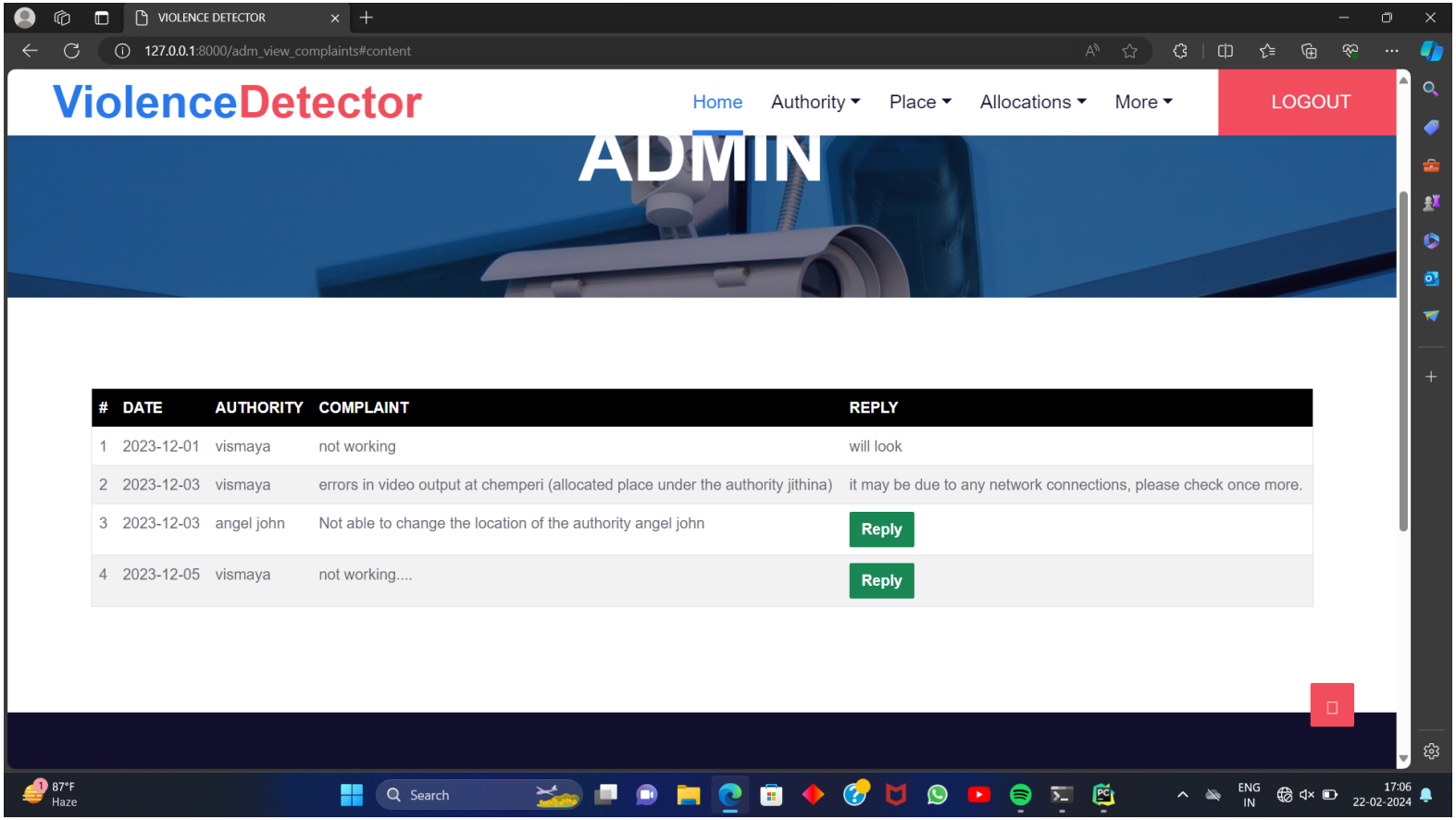


Figure: Complaints

IMPLEMENTATION

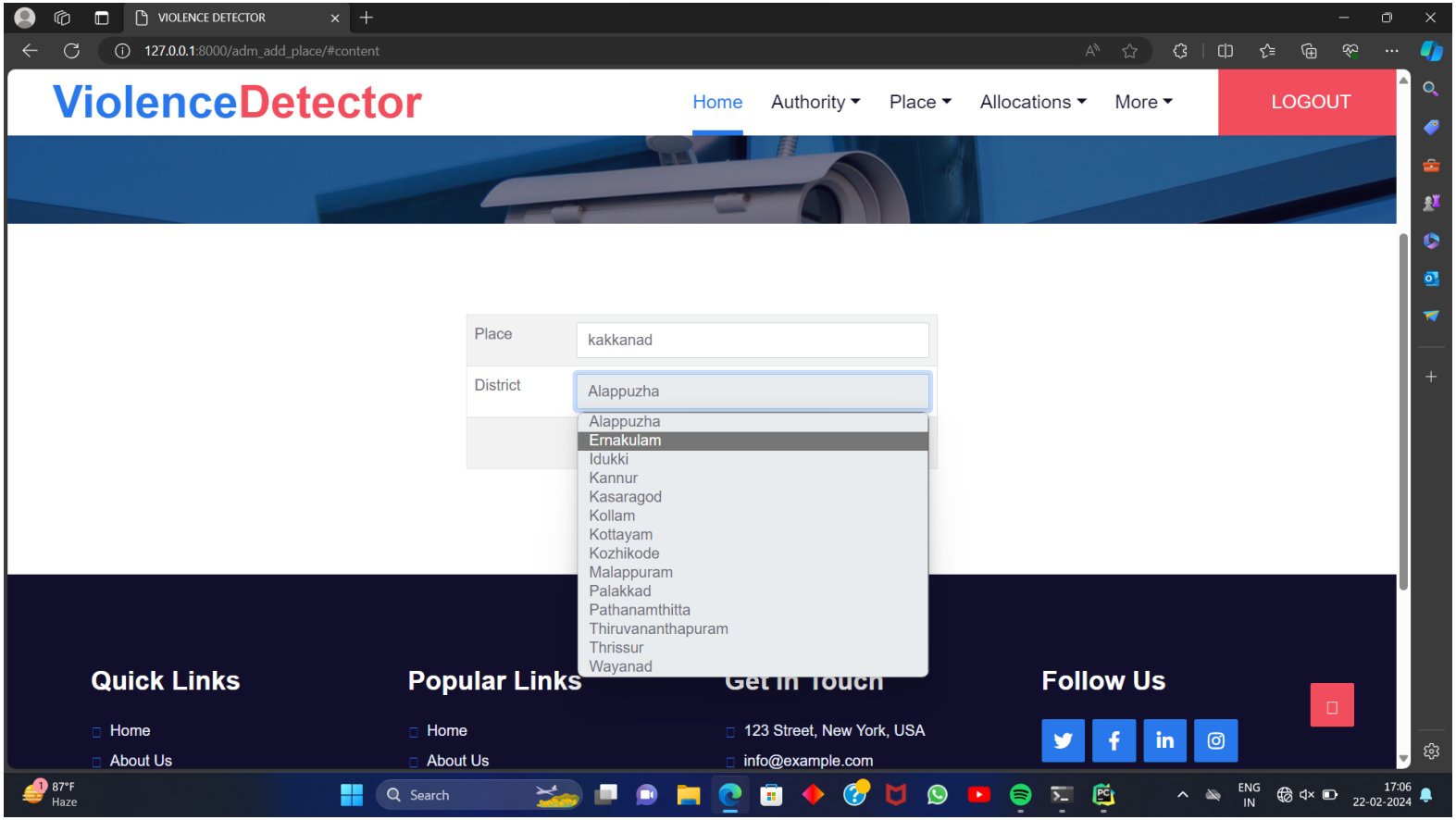


Figure: Place Allocation

IMPLEMENTATION

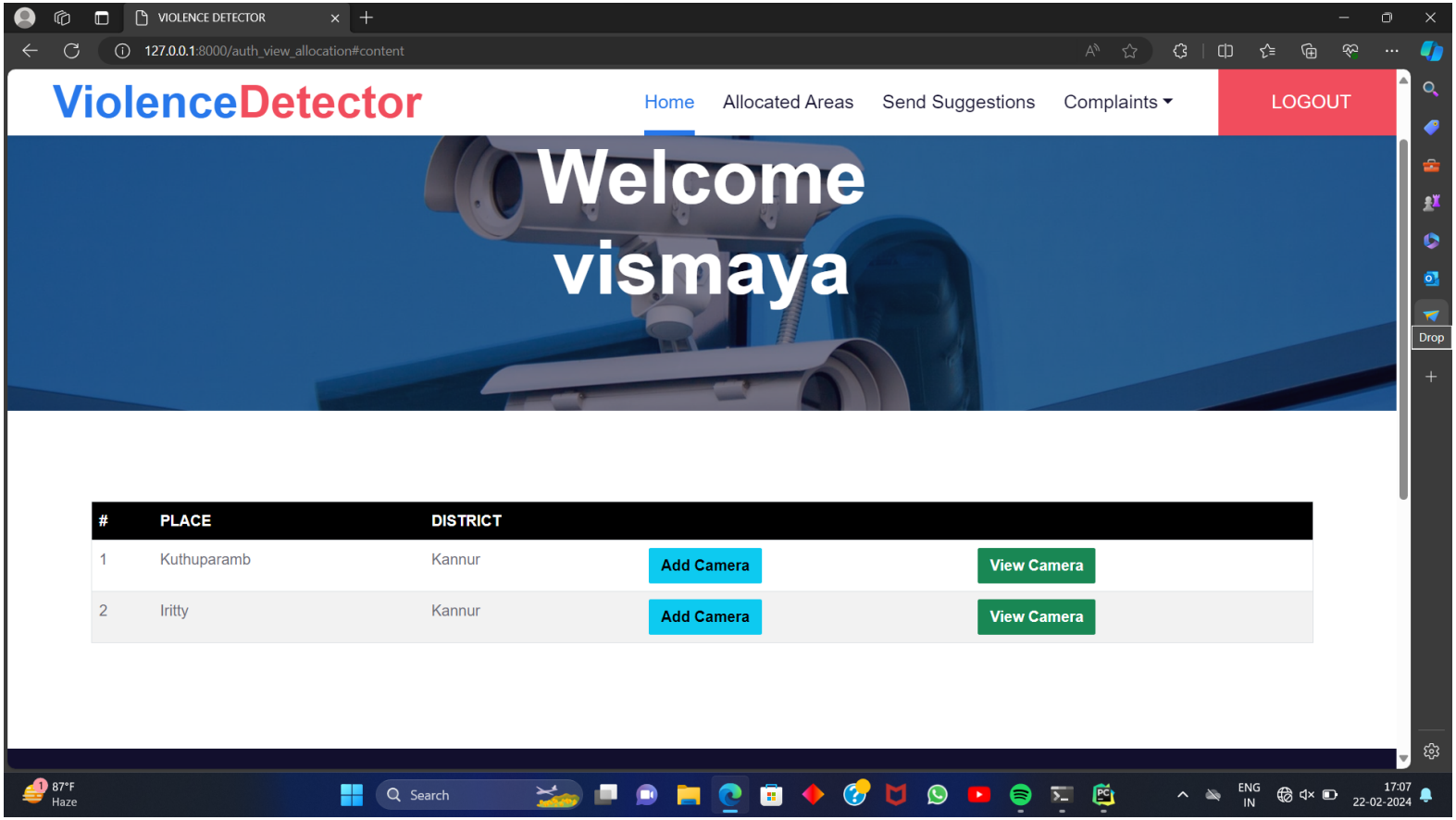


Figure: Alotted Cameras

IMPLEMENTATION

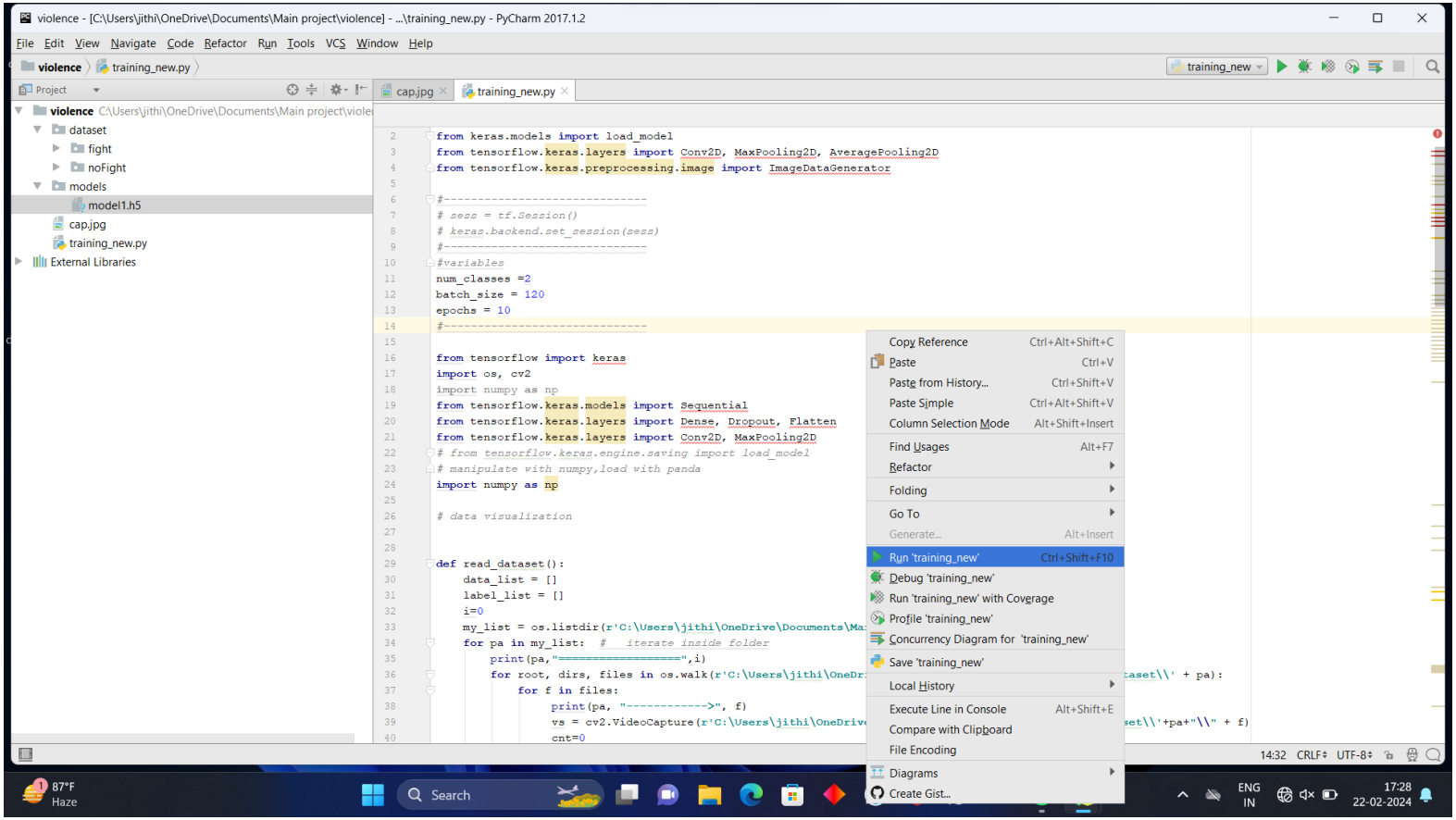


Figure: Code

IMPLEMENTATION

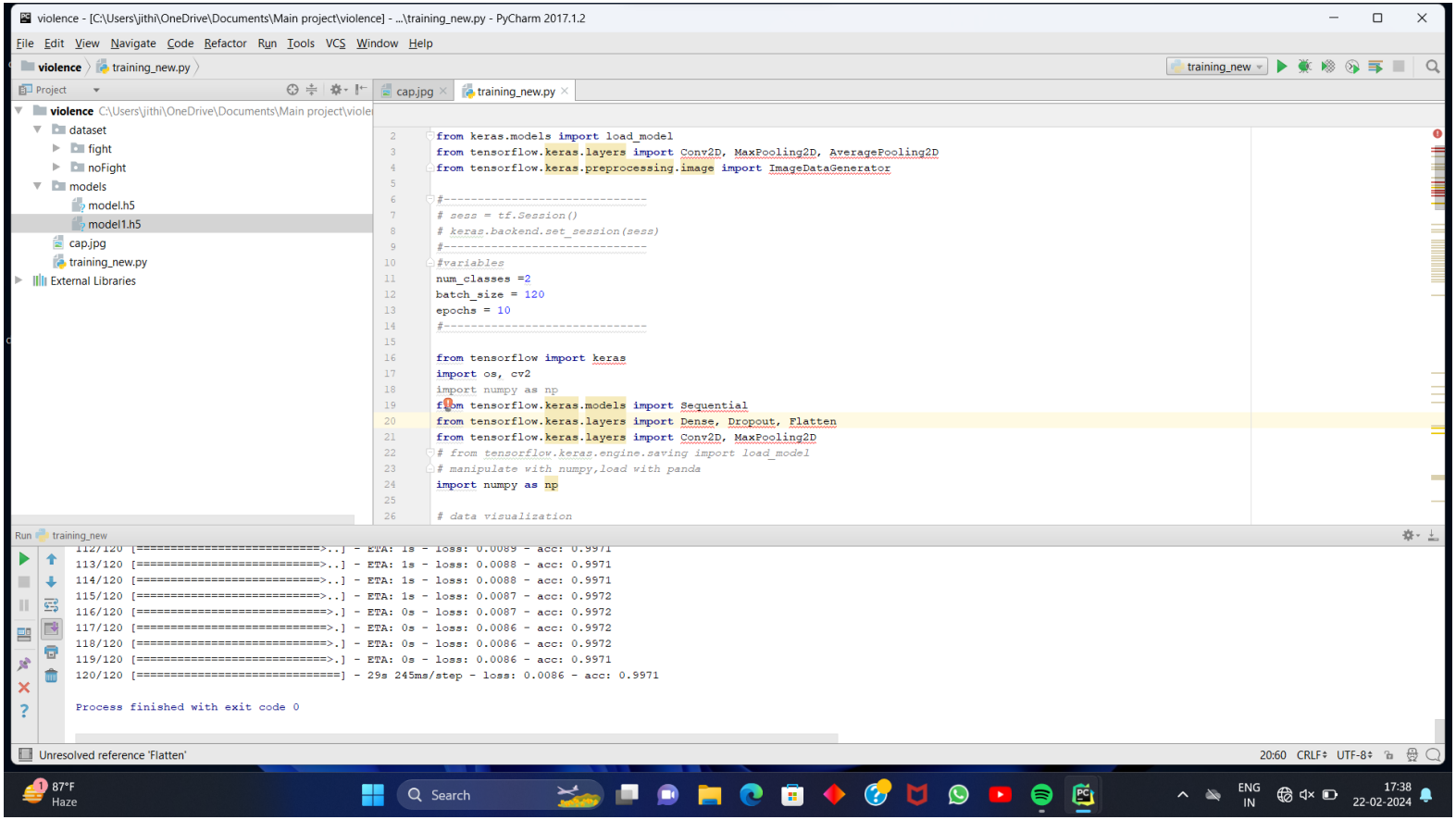


Figure: Training

IMPLEMENTATION

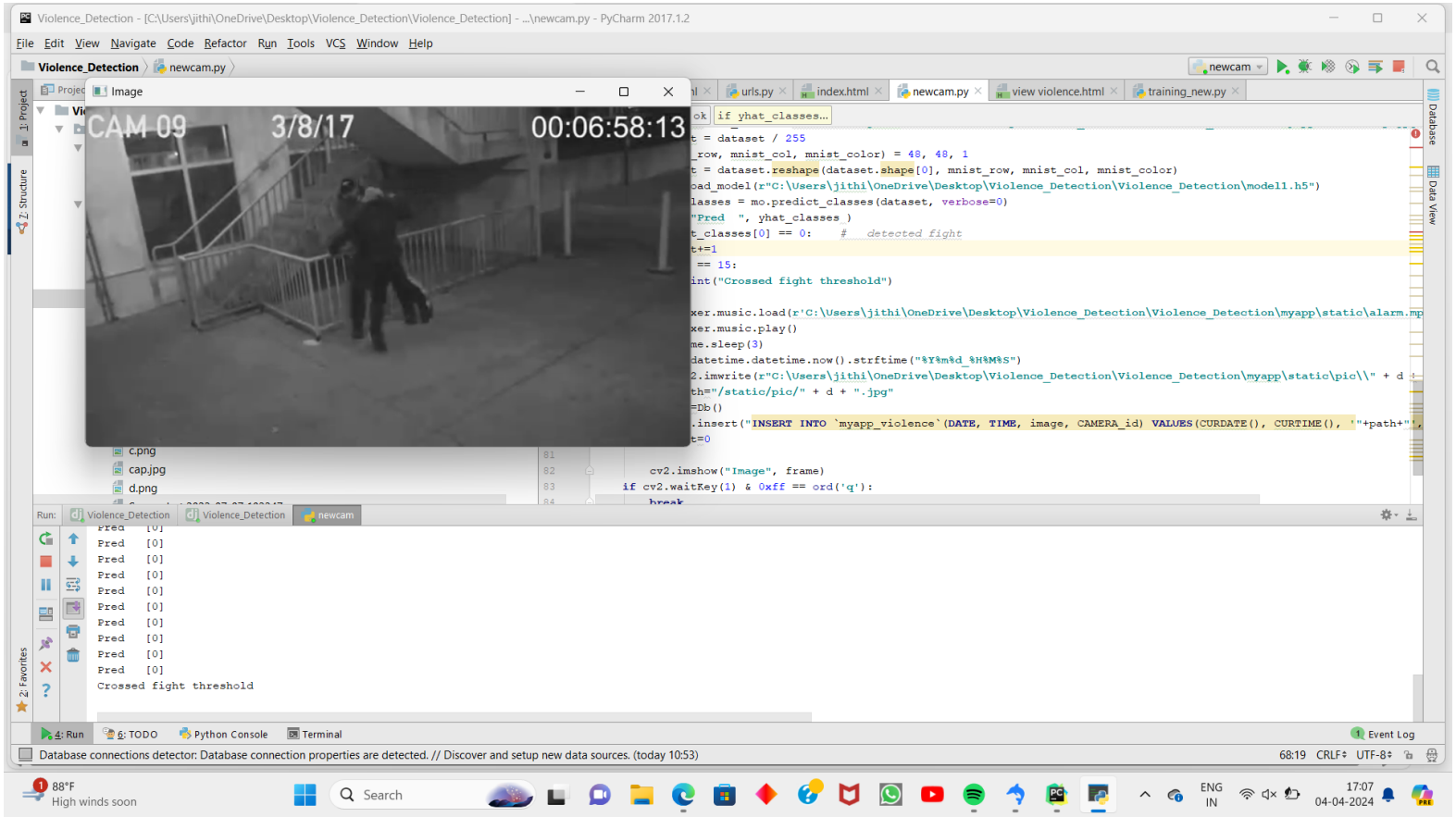


Figure: Violence Detection

IMPLEMENTATION

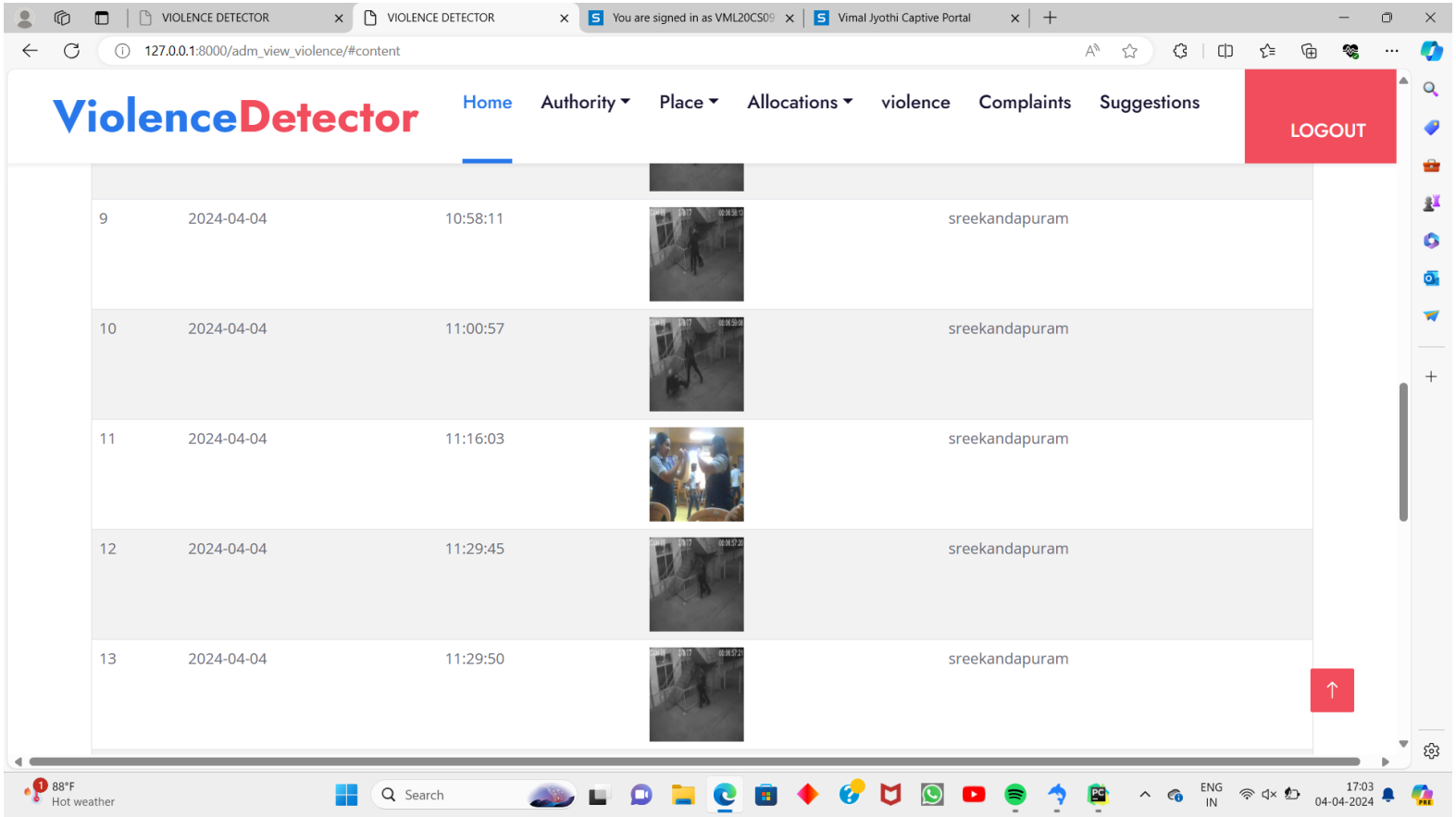


Figure: Alert Messages

RESULT

- The result of this project is accurately and efficiently detecting instances of violence in video streams.
- From the captured frames, violence is detected using the CNN algorithm and an alert is passed. The detected images and the corresponding location are viewed by the admin and authority.

CONCLUSION

- As society confronts evolving security challenges, the development of an efficient violence detection system for automated surveillance is of utmost importance.
- By combining deep learning and computer vision, this project strives to offer a promising solution to enhance security and safety in an increasingly complex world.

REFERENCES

[1] VIKTOR DÉNES HUSZÁR , VAMSI KIRAN ADHIKARLA , IMRE NÉGYESI , AND CSABA KRASZNAY,"Toward Fast and Accurate Violence Detection for Automated Video Surveillance Applications" ,IEEE Access, Journal Article, Volume 11 ,February 2023

[2] JOELTON CEZAR VIEIRA , ANDREZA SARTORI , STÉFANO FRIZZO STEFENON , FÁBIO LUIS PEREZ, GABRIEL SCHNEIDER DE JESUS , AND VALDERI REIS QUIETINHO LEITHARDT "Low-Cost CNN for Automatic Violence Recognition on Embedded System" ,IEEE Access, Journal Article, Volume 10 , 2022

[3] GHADAH ALDEHIM, MASHAEL M ASIRI, MOHAMMED ALJEBREEN, ABDULLAH MOHAMED, MOHAMMED ASSIRI , AND SARA SAADELDEEN IBRAHIM,"Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems",IEEE Access, Journal Article, Volume 11 , August 2023

[4] ABID MEHMOOD" Efficient Anomaly Detection in Crowd Videos Using Pre-Trained 2D Convolutional Neural Networks",IEEE Access, Journal Article, Volume 9 ,October 2021

REFERENCES

[5] H. Chen and W. Ye, "Classification of Human Activity Based on Radar Signal Using 1-D Convolutional Neural Network," in IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 7, pp. 1178-1182, July 2020

[6] J. Ha, J. Park, H. Kim, H. Park and J. Paik, "Violence detection for video surveillance system using irregular motion information," 2018 International Conference on Electronics, Information, and Communication (ICEIC), Honolulu, HI, USA, 2018

[7] C. Shripriya, J. Akshaya, R. Sowmya and M. Poonkodi, "Violence Detection System Using Resnet," 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2021

[8] A. -M. R. Abdali and R. F. Al-Tuma, "Robust Real-Time Violence Detection in Video Using CNN And LSTM," 2019 2nd Scientific Conference of Computer Sciences (SCCS), Baghdad, Iraq, 2019