Second Progress Report
on
Real Time Anomaly Detection in CCTV Surveillance

Submitted in the partial fulfilment of the Degree of Bachelor of Technology

(Computer Science and Engineering)

Submitted by

Avdhesh Chaudhary (02315002719)

Amar Sinha (04415002719)

Chirag Aggarwal (04615002719)

Under the supervision of

Mr. Vikrant Shokeen

(Assistant Professor CSE)

Department of Computer Science and Engineering

Maharaja Surajmal Institute of Technology Janakpuri, New Delhi.

2019-23

INDEX

Project Overview

Anomaly Detection System can be seen as a real time surveillance program designed to automatically detect and account for the signs of threatening activities immediately. We plan to use two Deep Learning models to detect and classify levels of high movement in a video frame. We plan to treat videos as segments and will define Anomalous(threatening) and Normal(safe) segments. From there, a detection alert can be raised in the case of a threat, indicating the suspicious activities at an instance of time. Further, we will recognize the following 12 anomalous activities - Abuse, Burglar, Explosion, Shooting, Fighting, Shoplifting, Road Accidents, Arson, Robbery, Stealing, Assault, and Vandalism. Detecting these anomalies would provide better security to the individuals.

To solve the above-mentioned problem, we will apply deep learning techniques used which would create phenomenal results in the detection of the activities and their categorization. Here, two Different Neural Networks: CNN and RNN are proposed.

Objective

The main objective of our project is to develop real time Anomaly detection in CCTV surveillance.

Objectives:

1. To reduce probability of error in anomaly detection in CCTV surveillance.
2. To reduce time of finding the video segment in which anomalous activities happen by making the process real time.
3. To increase accuracy of automatic threat detection in CCTV surveillance.
4. To increase the reliability of the system by making it more generalized and training it on 12 anomalous activities videos data.
5. To increase operational efficiency.

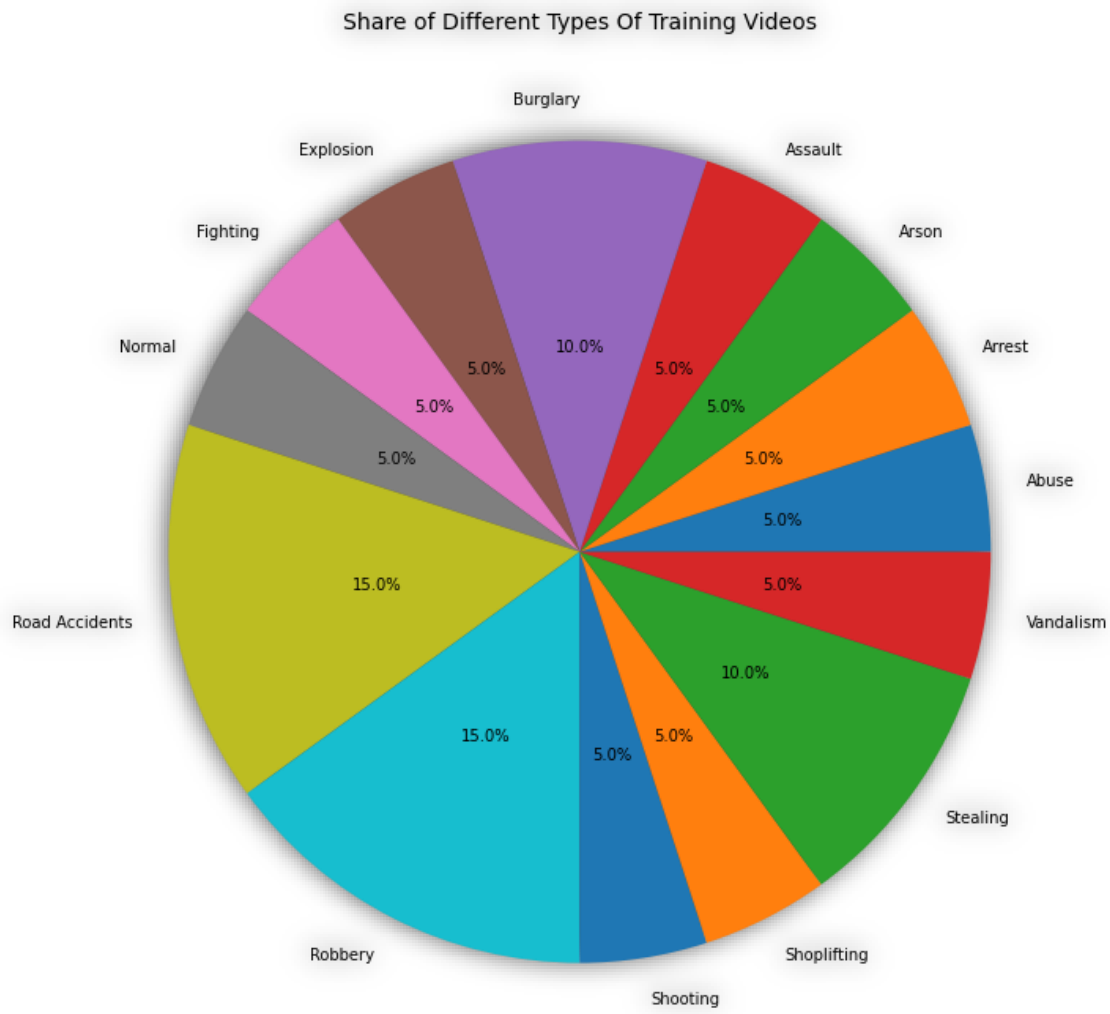
Objectives Achieved

The below Mentioned Objective has been achieved:

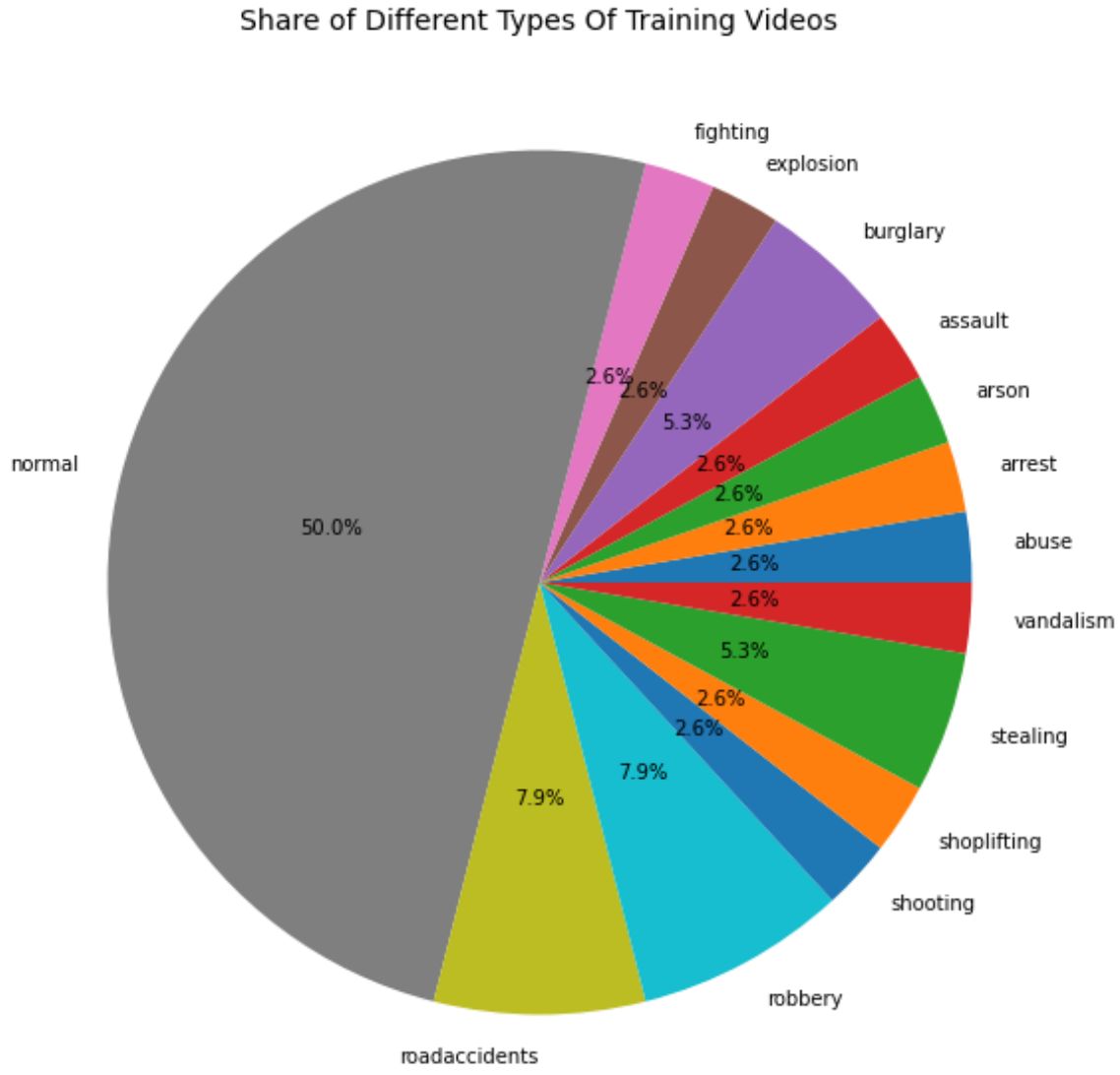
1. To reduce probability of error in anomaly detection in CCTV surveillance.
2. To increase operational efficiency.

The Following tasks are to make this objective completed:

1. **Dataset Fixing:** Earlier, the dataset had a greater proportion of anomalous videos as compared to normal videos. Such a dataset would have led to lots of "False Positive" and eventually resulting in lower precision, this was fixed by adding more normal videos to the dataset, restructuring the dataset and eventually making a customized train-test split for the same.



Earlier Dataset



Updated Dataset

2. **Transfer Learning and Feature Extraction:** Post conversion from videos to frames, each frame was fed to CNN where, the feature extraction network is typically a pretrained InceptionV3 architecture. The first subnetwork is a region proposal network trained to classify objects from the background, and the second subnetwork is trained to classify the detected objects' action alone (assault, arrest, abuse, normal etc.).

3. **CNN Model Working:** The output of the Inception model is passed to the input of the CNN which isn't the final classification model. Rather, the outcome of the last pooling layer is extracted which is a vector containing 2,048 features to feed as an input to RNN. The vector is referred to as a high-level feature map.

4. **Grouping of feature maps into a single pattern:** To give the framework a sense of the sequence, multiple prepossessed frames are considered. This chunk is then used to make the final classification. A chunk of these frames classifies a temporal segment of the video and can provide a sense of motion. For this, some feature maps are stored which are predicted by the inception model (CNN), generated in that fixed period of the video. Low-level features have been considered to generate a high-level feature map. These features are used for finding shapes and objects in computer images. This single combined feature map is then passed to the RNN. The reason to pass the feature map instead of the frame itself is to reduce the training complexity of the RNN.

Objectives In Progress

1. To increase accuracy of automatic threat detection in CCTV surveillance.
2. To create a demonstrable supporting project for the model.

References

1. Andrej Karpathy and Lei Fei-Fei. "*Deep Visual-Semantic Alignments for Generating Image Descriptions*", Department of Computer Science, Stanford University, 2015.
2. Vipin Shukla, Gaurav Kumar Singh and Dr. Pratik Shah. "*Automatic Alert of Security Threat through Video Surveillance System*", Department of Nuclear Engineering, Department of Electrical Engineering Pandit Deendayal Petroleum University, India, July 2013.
3. Utkarsh Contractor. "*Thread Detection in Surveillance Videos*", SPARK-AI Summit 2019.
4. Atif Jan and Gul Muhammad Khan. "*Real World malicious event recognition in CCTV recording*", January 2022.
5. Virender Singh, Swati Singh and Dr. Pooja Gupta. "*Real Time Anomaly Recognition Through CCTV Using Neural Networks*", 2020.
6. Abdul Jaleel, Syed Khaldoon Khurshid, Rehman Mustafa, Khalid Mehmood Aamir, Madeeha Tahir and Ahmad Ziar. "*Towards Proactive Surveillance through CCTV Cameras under Edge-Computing and Deep Learning*", March 2022.