Anomaly Detection in Video Surveillance (ADVS)

By

Capt Hamza Fiaz

Capt Ali Shaham

Capt Faheem Akbar

Capt Abdul Rafey

Supervised by:

Lt Col Khawir Mehmood

Submitted to the faculty of Department of Computer Software Engineering,
Military College of Signals, National University of Sciences and Technology, Islamabad,
in partial fulfillment for the requirements of B.E Degree in Software Engineering.

June 2022

In the name of ALLAH, the Most Benevolent, the Most Courteous

CERTIFICATE OF CORRECTNESS AND APPROVAL

This is to officially state that the thesis work contained in this report

"Anomaly Detection in Video Surveillance"

is carried out by

Capt Hamza Fiaz, Capt Ali Shaham, Capt Faheem Akbar, Capt Muhmmad Abdul Rafey

under my supervision and that in my judgement, it is fully ample, in scope and excellence, for the degree of Bachelor of Software Engineering in Military College of Signals, National University of Sciences and Technology (NUST), Islamabad.

Approved by

Supervisor
Lt Col Khawir Mehmood

Date: 31 May 2022

DECLARATION OF ORIGINALITY

We hereby declare that no portion of work presented in this thesis has been submitted in support of another award or qualification in either this institute or anywhere else. Reference material where used has been cited properly.

ACKNOWLEDGEMENTS

Allah Subhan'Wa'Tala is the sole guidance in all domains.

Our parents, colleagues, mentor Zain and most of all supervisor **Lt Col Khawir Mehmood**

without whose guidance this project would not have succeeded.

The group members, who through all adversities worked steadfastly.

COPYRIGHT

Plagiarism Certificate (Turnitin Report)

This thesis has **12%** similarity index. Turnitin report endorsed by Supervisor is attached.

_____

Capt Hamza Fiaz

00000280981

_____

Capt Faheem Akbar

00000281007

_____

Capt Ali Shaham

00000281015

_____

Capt Abdul Rafey

00000281013

_____

Supervisor

Lt Col Khawir Mehmood

Anomaly Detection in Video Surveillance

ORIGINALITY REPORT

12%	8%	4%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	4%
2	www.arxiv-vanity.com Internet Source	2%
3	Submitted to University of Hull Student Paper	1%
4	tesis.ipn.mx Internet Source	<1%
5	shura.shu.ac.uk Internet Source	<1%
6	Virender Singh, Swati Singh, Pooja Gupta. "Real-Time Anomaly Recognition Through CCTV Using Neural Networks", Procedia Computer Science, 2020 Publication	<1%
7	Waqas Sultani, Chen Chen, Mubarak Shah. "Real-World Anomaly Detection in Surveillance Videos", 2018 IEEE/CVF	<1%

ABSTRACT

In this project, we have developed an anomaly detection system which makes use of Machine Learning to detect anomalies to include Violence, Theft, Accident, Arson, and Abuse. This would be accomplished by using deep neural networks. Two approaches have been adopted to fulfil the requirement, Multiple Instance Learning approach and MobilenetV2 approach. Both use convolutional neural networks to train machine learning models on different datasets. Both approaches have been implemented on different platforms with different frameworks. Predesigned datasets have been used for training models, datasets comprise of large number of videos containing normal and anomalous behaviors. After training on different frameworks, models are designed to detect anomalies from real world scenarios. Both models have been rigorously tested and display high accuracy for the anomalies mentioned above. Model is then deployed on an interface which takes the video as an input and displays results, either as a graph or in the form of video depiction as per the requirement, for our different approaches. The output extracted can further be utilized for deployment on the end system for real-time anomaly detection in surveillance videos.

Table of Contents

List of Figures

Chapter 1: Introduction

Constant monitoring of the real time surveillance videos to detect any sensitive situation requires manpower and monitoring resources. It is tiresome and hectic if done over a long period of time. With the advancements in technology and especially in the field of Artificial Intelligence, it has become possible to automate such systems. Anomaly Detection in Video Surveillance aims to achieve this with maximum accuracy and efficiency. In this system, automation of the surveillance system would be achieved for various anomalies using computer vision algorithms to notify in case of a set of predefined possible anomalies such as Violence, Accidents, Explosion, Theft, Abuse and Arson. ADVS will be able to identify catch these anomalies, notify users, ring alarms if necessary and ensure maximum automation of the surveillance systems.

1.1 Overview

Surveillance cameras are becoming vital in modern day scenario because of the overwhelmingly heinous nature of crimes being committed on day-to-day basis. Keeping pace with monitoring of these cameras and making utilization to the full extent has unfortunately not been possible. This has led to an increase in crimes and confidence gaining of these criminals. Another aspect of these cameras is there is a huge ratio deficit between cameras and human monitoring operators. Therefore, it is not humanly possible to monitor all of them with accuracy and efficiency plus there is also an element of human error which reduces the already insufficient efficiency of human monitoring operators**.**

1.2 Problem Statement

Real world anomalies are diverse and complex, and the autonomous system dictates that the system works with minimum to no supervision. Anomalies occur at all times and any behavior that does not correspond to normal routine is considered to be an anomaly. Therefore, there is a requirement of a system that specifically enlists all anomalies autonomously highlighting the exact timeframe of these anomalies. All these elements combined, create a vacuum for developing an intelligent system for improved security.

1.3 Proposed Solution

The proposed solution would comprise of state-of-the-art algorithms for anomalous behavior detection autonomously. These anomalies in question would be defined exactly that is **Arson, Accident, Explosion, Theft, Abuse** and **Violence**. The project has been developed in an iterative form. Firstly, it would be developed to detect anomalies on videos and in subsequent stage it would be implemented on surveillance cameras.

We have followed two different approaches to train our machine learning models for detecting anomalous activities. The variation in both approaches lies in the methodology to train the models. Our initial approach has been the primary approach for this project, which uses MIL (Multiple Instance Learning) to train our model while the secondary approach uses Mobilenetv2, which is a model used for image classification. Both these approaches make use of different datasets to train their models, which are then used to detect anomalous behavior in the videos. Both approaches will be discussed in detail in the following chapters of this document. In each section,

a brief explanation has been provided for both approaches. We will be identifying these approaches as the MIL approach and the Mobilenetv2 approach.

1.3.1 MIL Approach

It uses Multiple Instance Learning to create bags for anomalous and normal videos to follow the training of the model. Once the model is trained it has the ability to identify and capture the exact frames where an anomaly starts and terminates. The graph in Figure 1 indicates the working. The system has been rigorously tested and trained on different real world anomalous scenarios. Upon the detection of anomaly, the system detects the anomalous behavior and frames are captured.

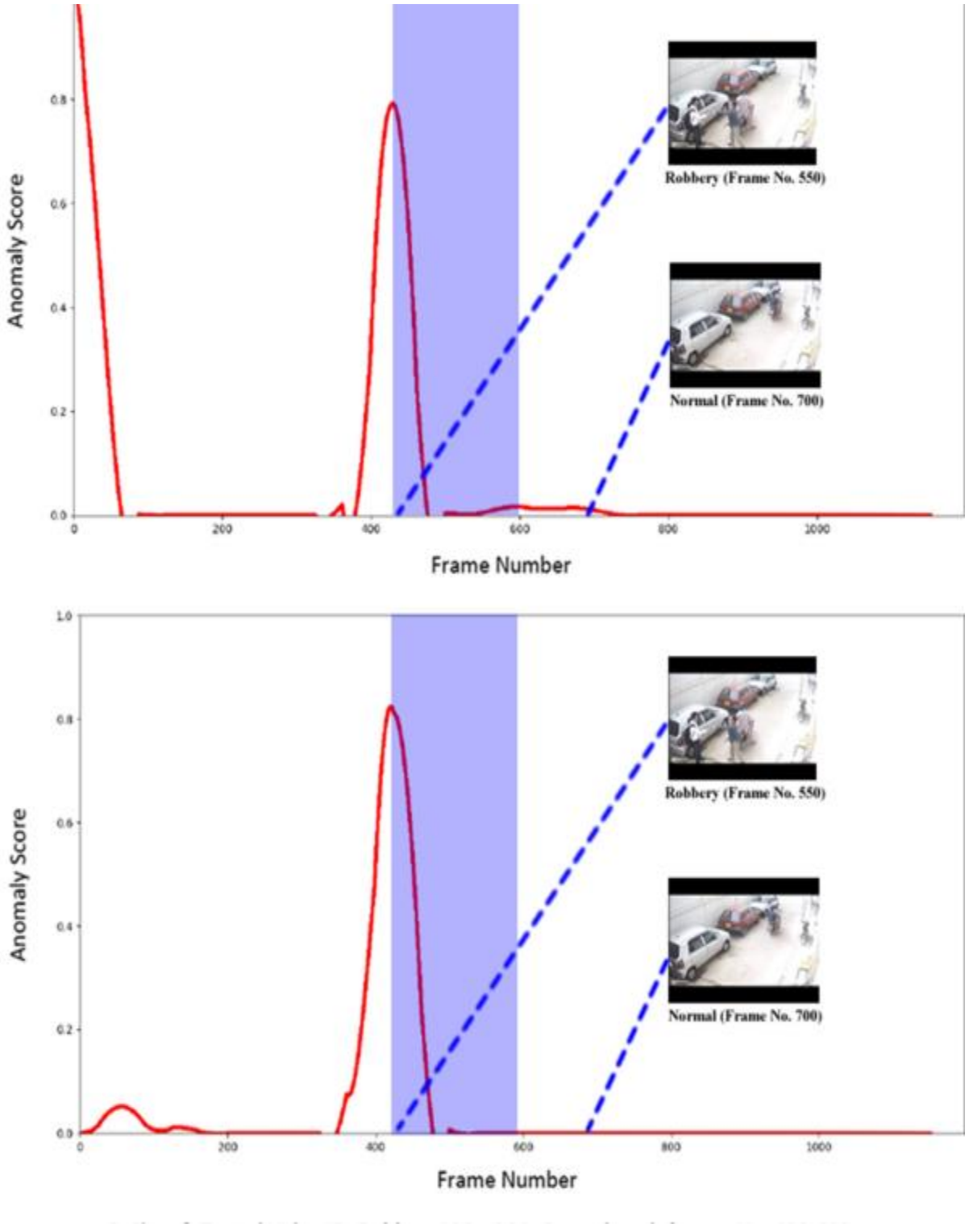

Fig 1 Anomaly Detection at frame 380-460

1.3.2 MobilenetV2 Approach

MobilenetV2 is an image classification model. It is suitable for mobile devices or devices with low computational power. MobilenetV2 achieves the training of the model with the help of a set of convolution blocks which are the layers of this function to achieve the maximum accuracy in any image classification problem. After the model is trained on the dataset, it is able to analyze videos by cutting them into frames and then classifying them as anomalous or normal. We restrict this approach to detect the violent activities using the MobilenetV2 function on a real-life violence situations dataset.

1.4 Working Principle

The project primarily works on the principles of image processing amalgamated with machine learning algorithms. The project is divided into different modulus and every module is inter-woven with the next module. The list of modules is as under:

- Dataset
- Model training and testing
- Implementation
- **Detection of Anomaly**

1.4.1 Dataset:

The primary part of project is the preparation of dataset. For rigorous training a massive dataset is the requirement. Owing to limitations of majority of datasets, a new requirement emerged requiring an extensive dataset with real world true events, in which some events contain anomalies. Preexisting datasets have been selected owing to their massive scales.

1.4.1.1 MIL Approach:

UCF Crime has been used in our MIL approach. This dataset is University of Central Florida crime dataset. It comprises of 1900 real world surveillance videos of more than 128 hours comprising of both normal and anomalous videos. The dataset is massive and one of its kind so far. Encompassing various type of anomalies out of which few anomalies have been chosen for this project to include arson, violence, explosion, accidents and shoplifting.

1.4.1.2 MobilenetV2:

Real-life Violence Situations Dataset has been chosen for MobilenetV2 approach. This is a large dataset containing 1000 violent videos from real-life situations and 1000 non-violent or normal videos. Violent videos are mostly real street fights taken from different environments and conditions.

1.4.2 Dataset training and processing:

1.4.2.1 MIL Approach:

The prepared dataset is used as input to train anomaly detection model using machine learning. The dataset has been divided into two parts. The training part consists of 800 normal and 810 anomalous videos whilst the testing part consists of remaining 150 normal and 140 anomalous videos. With training iterations, the accuracy of model is improved figure 3 shows the improvement in model with iterations.

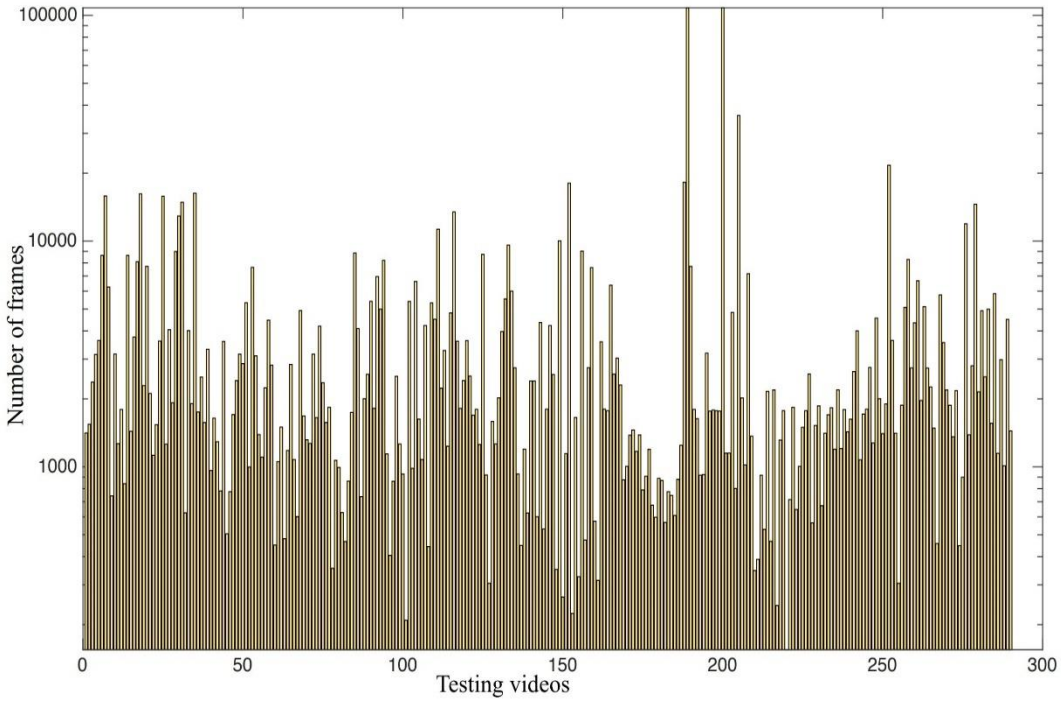

Fig 2 Distribution of Video Frames in Testing Set

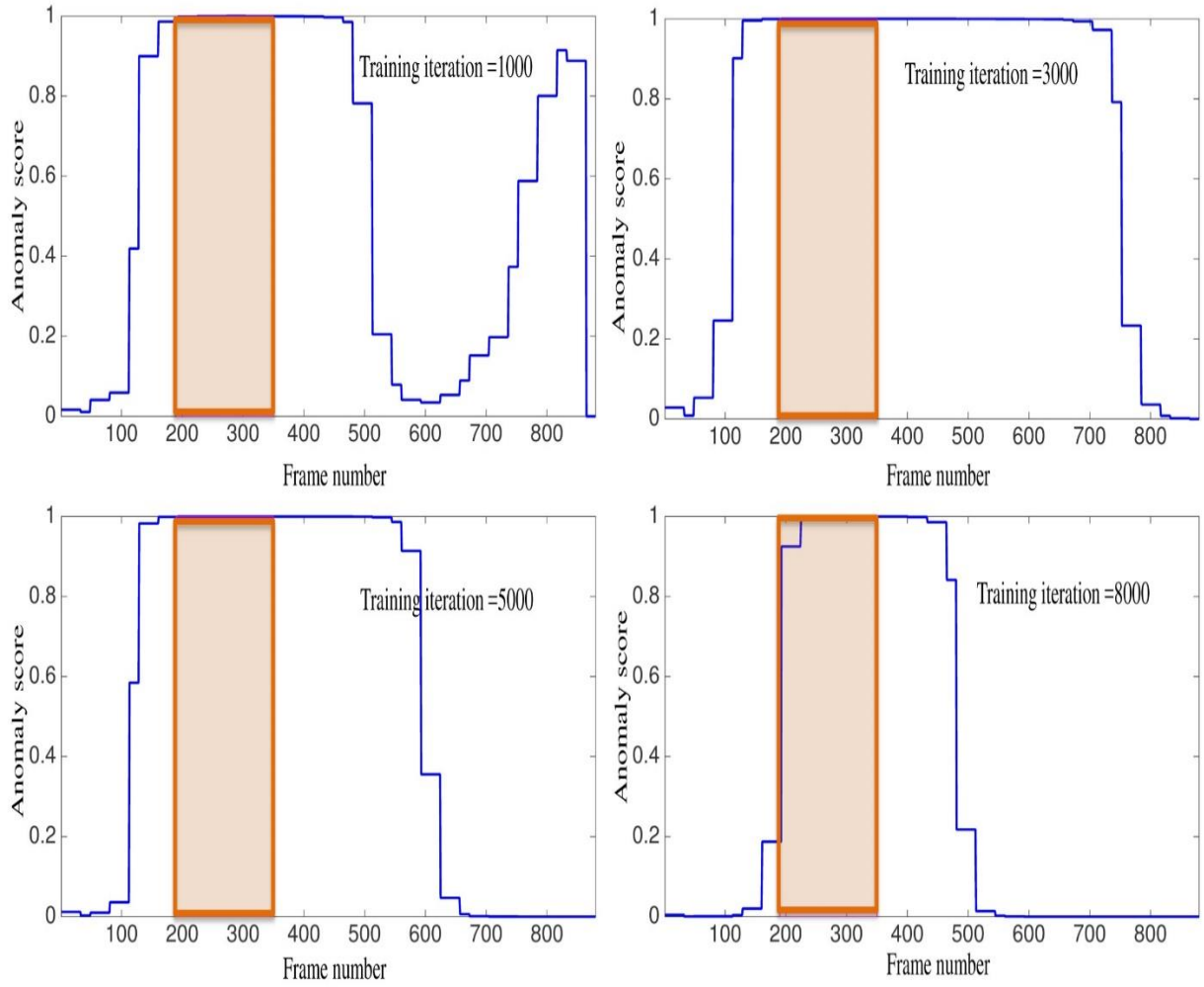

Fig 3 Training Iterations

1.4.2.2 MobilenetV2:

Real-life Violence Situations dataset is used as input to train anomaly detection model using machine learning. The model is trained on violent and non-violent videos from the dataset. After training the model using the MobilenetV2 function, we test the model on the same videos and get very high accuracy. We use 350 violent videos and 350 non-violent videos to train our model. Figure 4 shows the results of our testing on the dataset after extensive training of the model using 50 epochs.

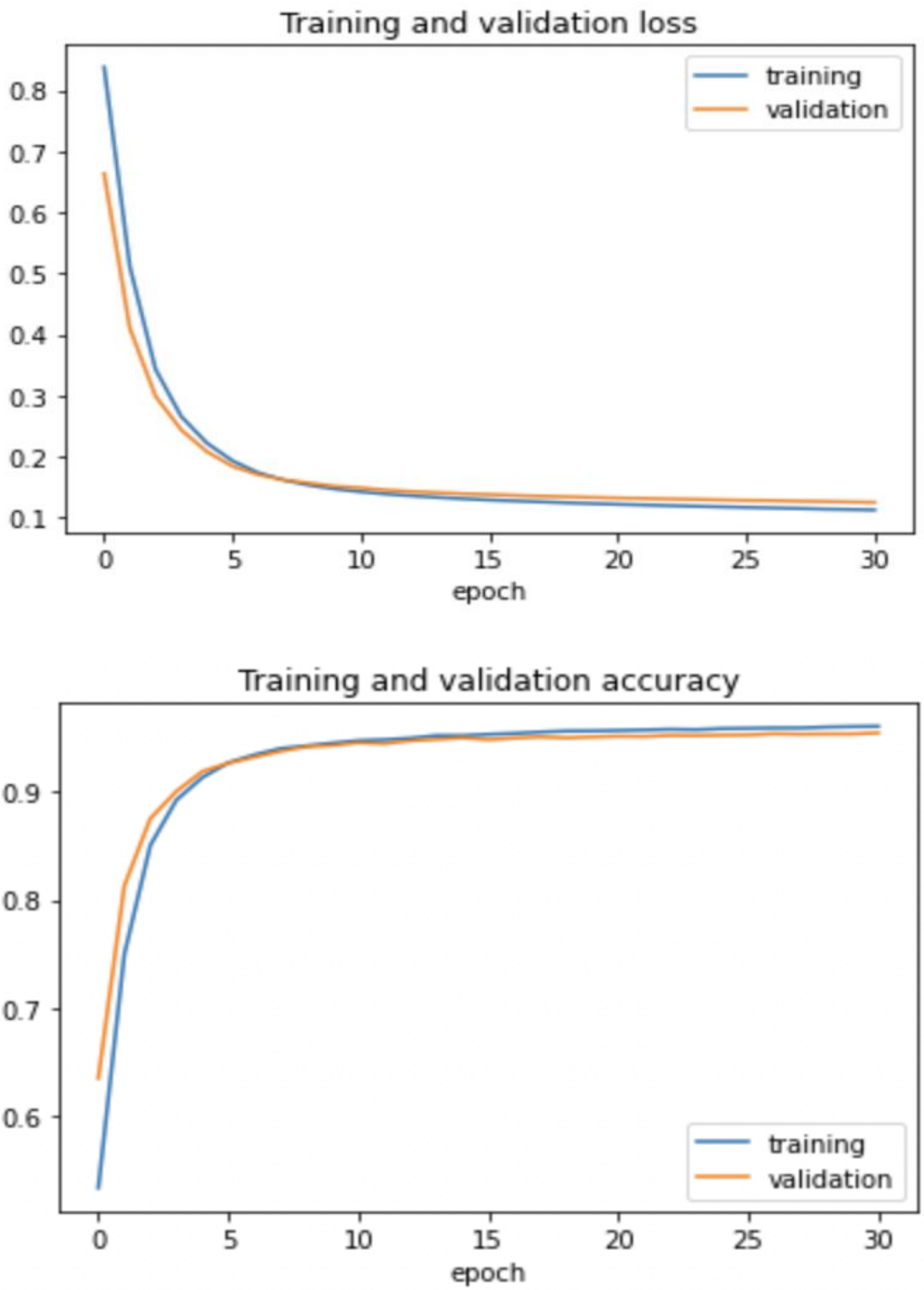

Fig 4 Training and Validation Loss and Accuracy

Figure 5 shows the correct and incorrect predictions. The top left are the videos which were predicted true negatives and bottom right as true positives, which amount to 96 percent of the videos in the dataset showing the high accuracy of our model after an extensive training with the MobilenetV2 function.

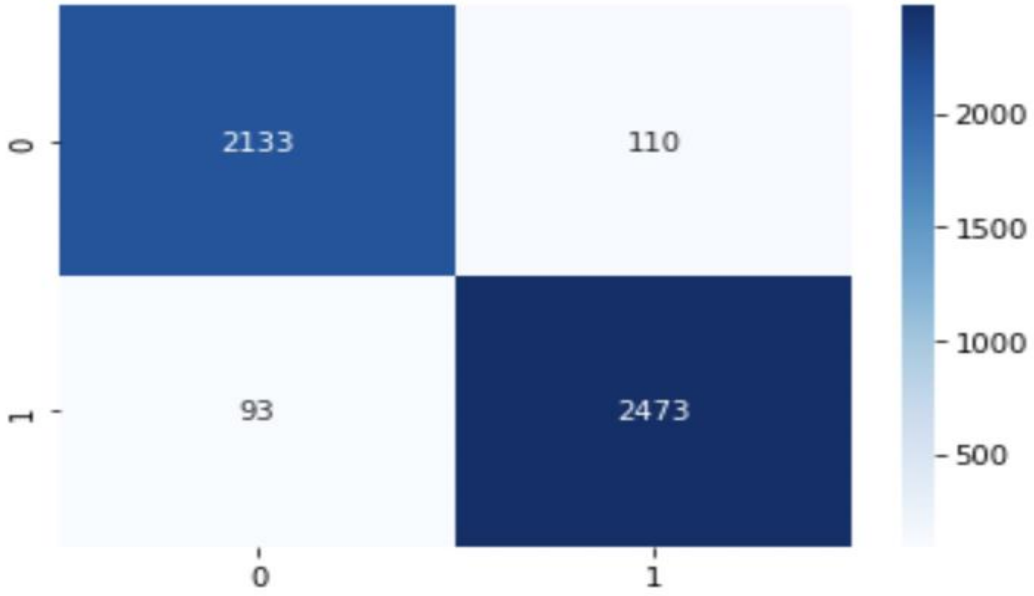

Fig 5 Correct and Incorrect Prediction

1.4.3 Functionality in Both User Interface:

1.4.3.1 MIL Approach:

MIL approach is designing, developing and implementing a separate anomaly detection system namely Anomaly Detection in Video Surveillance which detects other mentioned anomalies **Arson, Accident, Explosion, Shooting** and **Theft**. It detects **Normal** videos as well. The system is capable of extracting features from video, performing calculations on these features using C3D v1.0 and then applying trained model onto the extracted features to detect the beginning and

termination of anomaly at the frames that it occurred. Each video is segregated into 32 segments for feature extraction using C3D feature extraction tool. The implementation is tested using Keras, Theano, Python and PyQt5 modules. Implementation would be explained elaborately in chapter 4 Implementation and Code Analysis. Figure 8 shows the graph which is displayed as an output to show the anomaly detection in the. In this graph we can observe that in the initial 1000 frames there is a very high probability that an anomaly has occurred then after a brief pause from again an anomaly occurs with comparatively less probability and over a very brief period, then the graph depicts the normalcy detecting no anomaly and showing absolutely no probability of an anomaly for the rest of the video. This is just one of such representation of detecting anomalous behavior in a video.

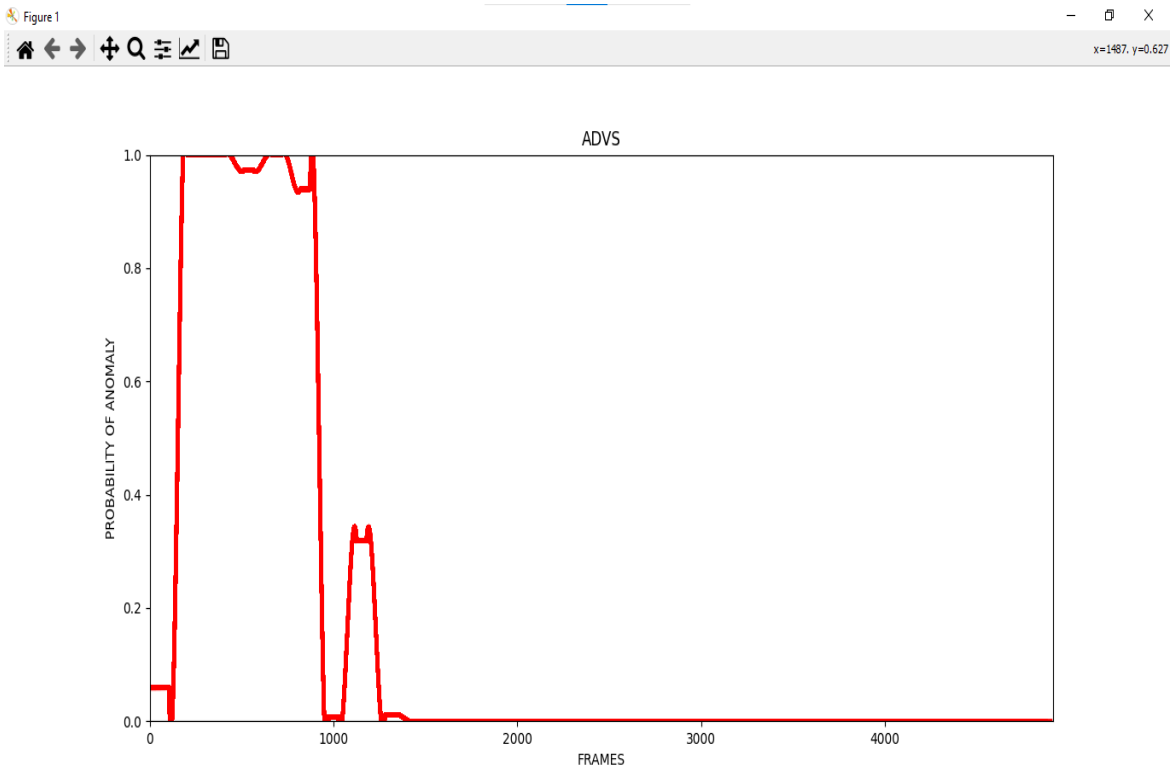

Fig 6 GUI MIL Approach

1.4.3.2 MobilenetV2 Approach:

MobilenetV2 approach, after training and testing the model with MobilenetV2 function, the model is saved and then loaded for analyzing videos for any violence, the model takes frames from video and displays output frames, each frame either displays violence as true or false. Throughout the video if there are scenarios where violence is taking place, the output frames show a message on the top left as shown in Figure 7 depicting violence as true and where the video has no violence those frames are returned with message as false for violence. In this way whole video is broken into frames and displayed as output. An output video file is also saved in the drive showing violence detection message in the complete video, turning red and displaying message ‘true’ for violence or green and displaying message ‘false’ for no violence as shown in the Figure 7.

Fig 7 GUI MobilenetV2 Approach

1.4.4 Detection of Anomaly:

1.4.4.1 MIL Approach:

The extracted outputs, Indicate the detection of anomaly from the exact frame number and the graph indicates the beginning and termination of the anomaly. If no anomaly is detected the graph remains stagnated and a straight line is painted on x-axis indicating no anomaly as highlighted in fig 5 above.

1.4.4.2 MoblenetV2 Approach:

The extracted output frames display the anomaly message as true or false. An output video is also created to analyze the anomaly detected throughout the video.

1.5 Objectives

1.5.1 Objectives:

"To design and create an innovative solution for detection of real-world anomalies that affect the public safety using machine learning and automate the system of monitoring surveillance cameras, thereby eliminating the margin and element of human error and also increasing efficiency and efficacy of surveillance systems, reducing administrative burden on human monitoring."

1.5.2 Academic Objectives:

- Development of an innovative Anomaly Detection in Surveillance Videos system
- To implement Machine Learning techniques and simulate the results in form of graphs for comprehensive understanding and subsequent utilization
- To reduce administrative load on human monitoring operators
- To utilize the surveillance systems to their full potential

1.6 Scope

This project finds its scope wherever there is a camera, or a surveillance system. This includes all sensitive installations in the modern world. Every institute these days is equipped with CCTV cameras and surveillance has become a necessity. To keep the environments secure, constant monitoring is required. This entails schools, colleges, universities, hospitals, Roads, Railway Stations, Offices, Airports and other Sensitive Installations. In short, all the installations with a monitoring system installed. The project is implemented iteratively, in first step it is implemented by uploading videos and detecting the anomalies. In future the project would be extended to real time anomaly detection from surveillance.

1.7 Deliverables

1.7.1 Anomaly Detection System

Anomaly detection system produces a graph as an output that illustrates the working and mechanism or in another case display the anomaly message on the video after processing. Elaborate details are presented in chapter 4 Implementation and Code Analysis.

1.8 Structure of Thesis

Chapter 2 Literature Review and the Background Analysis.

Chapter 3 Design and Architecture.

Chapter 4 Implementation and Code Analysis.

Chapter 5 Future Work

Chapter 6 Conclusion

Chapter 2: Literature Review

New product is introduced by modifying and enhancing the features of previously existing similar products. Literature review is an imperative step for development of an idea to a brand-new product. Likewise, for the development of anomaly system, a detailed study regarding similar projects is mandatory. Our research is divided as follows.

- Industrial Background
- Existing solutions and their drawbacks

2.1 Industrial background

Monitoring systems have been there since the invention of cameras. Various approaches have been followed in detecting anomalies. These approaches depended upon the very definition of an anomaly. Some argued that any deviation to normal behavior was considered to be an anomaly. Others implemented a goal-based approach where a specific behavior was considered an anomaly like fire or theft.

CCTV monitoring has always been a hassle and considered to be an extremely difficult task to automate because of the various factors involved. Cameras are constantly recording, and a requirement of instant anomaly detection would save lives. Sometimes an event is also missed owing to the paucity of observers (humans). A meaningful autonomous anomaly detection system is the need of the hour.

2.2 Existing solutions and their drawbacks

Two state-of-art approaches for anomaly detection would be compared with ours.

2.2.1 *Lu et el* Approach

Lu et al proposed a dictionary-based approach to identify and learn normal behaviors and using reconstruction errors to detect anomalies. Following this approach 7000 cuboids were extracted from each normal video and gradient based features in each volume were computed. After reduction of dimensions dictionary is learned using sparse representation. The accuracy of model was 65%

2.2.2 *Hassan et el* Approach

Hasan et al proposed a full convolutional feed forward deep auto-encoder based approach to identify local features and classifier. Using this implementation, network is trained on normal videos using the temporal window of 40 frames. After this reconstruction error is used to identify anomaly. The accuracy stood at 50%

2.2.3 *Our* Approach

Our approach has a deep neural network that has convolutional, max pool and fully connected layers. The approach allows us to extract features after necessary computation on videos. The approach is elaborately explained in chapter 3.1 Architecture. Our proposed model has achieved accuracy far more than both standard approaches mentioned above. The accuracy stands at **75.61%** for MIL approach as anomaly detection in videos is considered to be an extremely difficult challenge in machine learning. The accuracy stands at **96%** for MobilenetV2 approach in detecting violence.

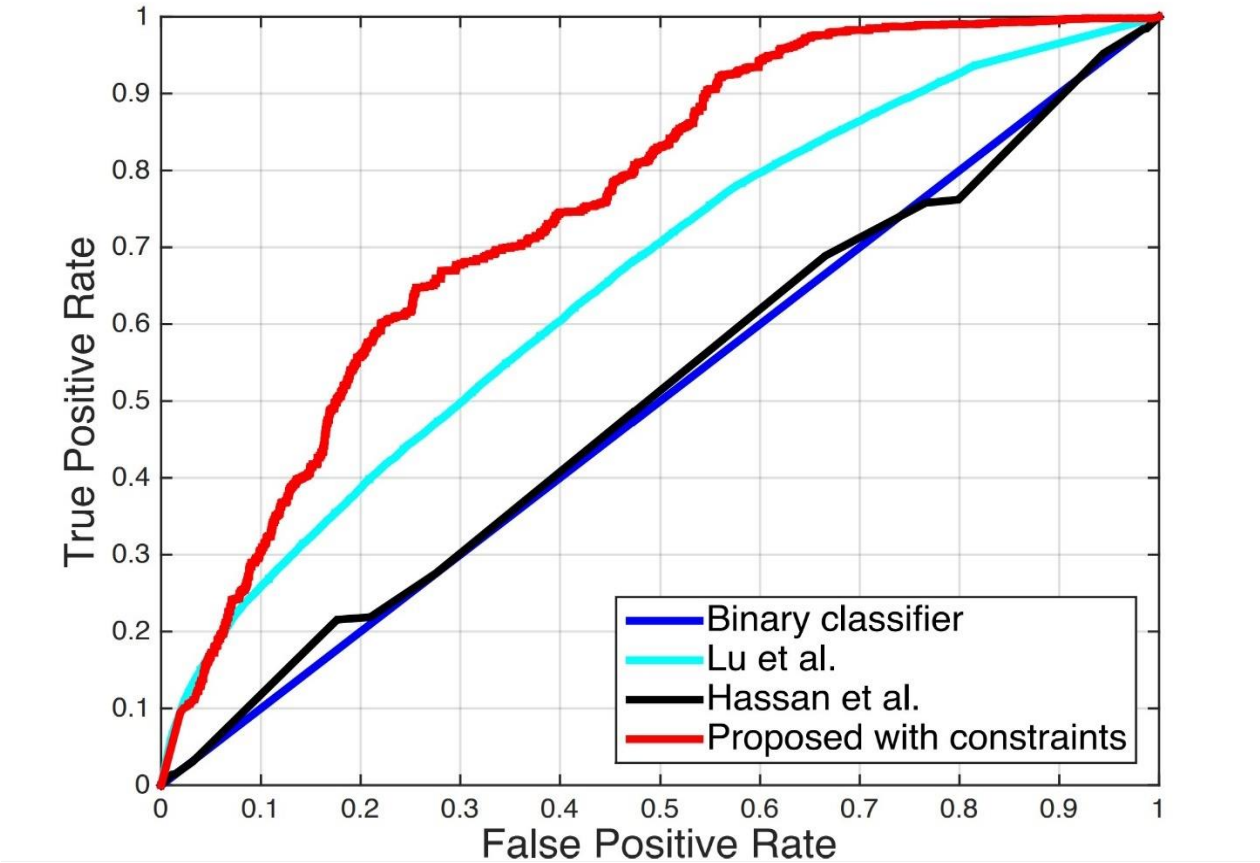

Fig 8 Receiver Operating Characteristics Graph

Chapter 3: Design and Architecture

3.1 Design

The design aims to achieve maximum accuracy in the anomaly detection, however in recent works on the same subjects, different designs have been used from implementation as mentioned in 2.2 Existing solutions and their drawbacks. In real-time systems, the challenge is to detect the anomalies in lightning speeds and notify users. In the various designs used, deep learning has resulted best for image classification and is found suitable for video classification as well. Apart from this there are certain other computer vision and digital image processing methods to achieve the anomaly detections in - real time.

Deep neural networks have been chosen for our approach, Model containing convolutional neural networks and recurrent neural networks to solve the problem. Our first neural network which is the convolutional network, reduces the intricacies in the model making it less complex and converting low level features to high level features, further the recurrent network will be classifying the features and returning the overall accuracy of the model, using the test and train data.

The design is implemented using various modules and dependencies which are mentioned in detail in chapter 4 Implementation and Code Analysis.

3.2 Architecture

Architecture of ADVS is explained thoroughly through a diagram that explains each step and its working. There are two sets of videos namely anomalous and normal videos. Each video is divided into equal 32 segments. Two bags of videos are formed as a result i.e., positive bag and negative bag of videos. Each video is then passed through C3D tool converting video into features

and those features are extracted. C3D feature extraction has a deep neural network that extracts the features. The layers contain a Convolutional layer→ Max Pool Layer→ Convolutional Layer→Max Pool Layer→Fully Connected Layer. This neural network adjusts weights and extracts instances that are transferred to either positive bag or negative bag. The highest instance score in each bag is the utilized as a trained model against which testing videos are weighed. The complete architecture is depicted in fig 9 Architecture of ADVS.

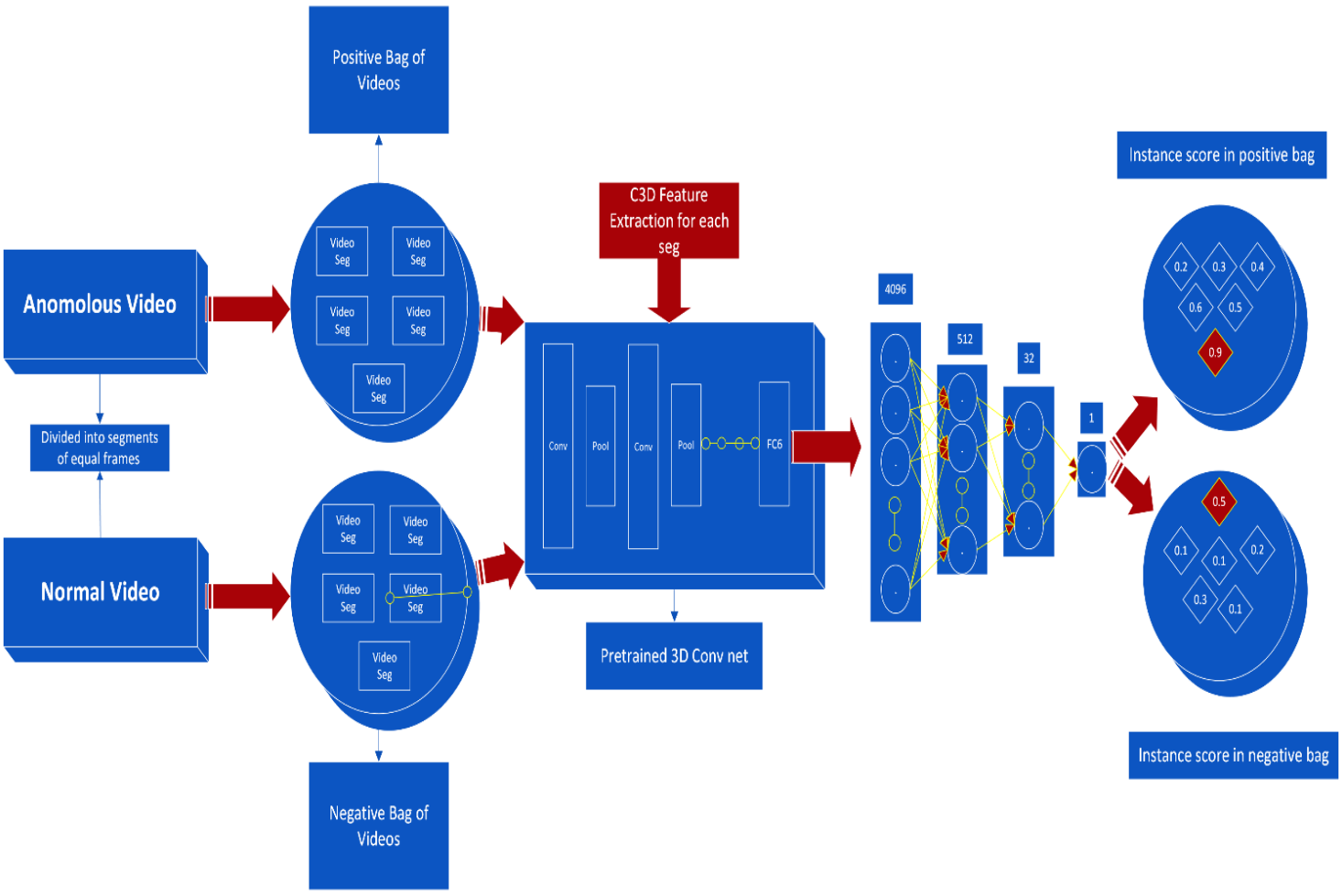

Fig 9 Architecture of ADVS

For the MobilenetV2 approach, we used a simpler architecture where we use the MobilenetV2 function to train and test the model. Figure 10 shows a visual representation of the functioning of MobilenetV2 function. We utilized multiple convolution layers where we used 1x1 convolutions and depth wise convolutions followed by Relu activation in the MobilenetV2 function to train the model.

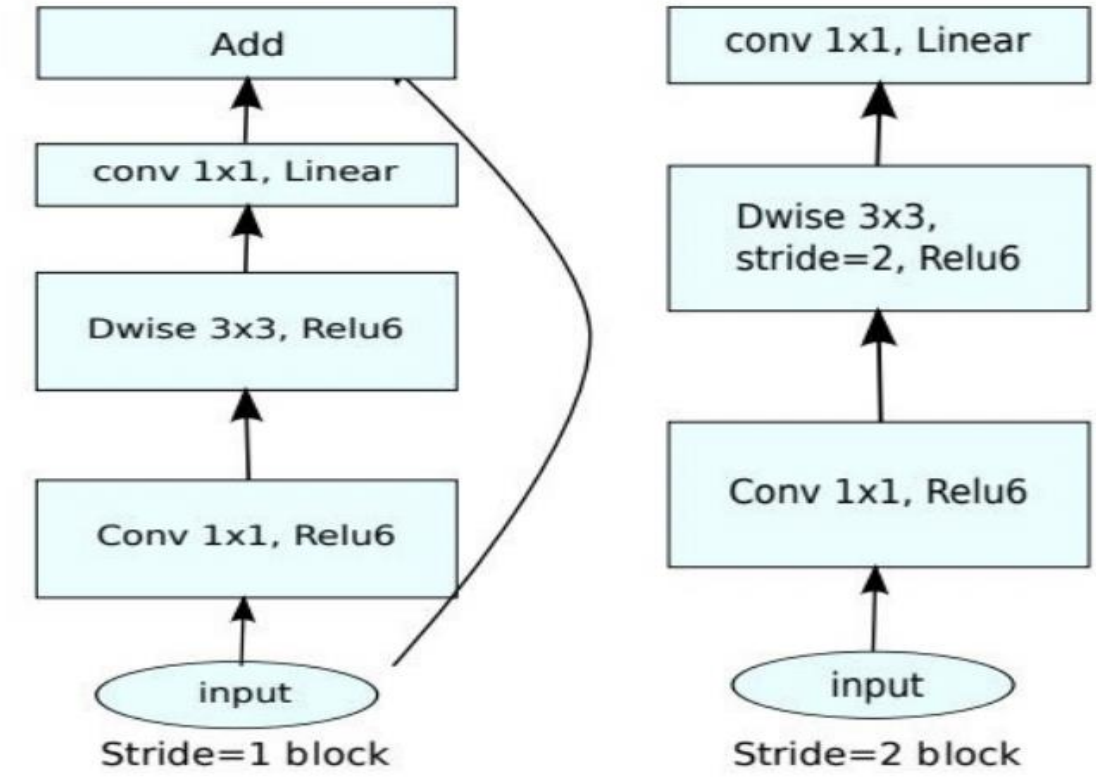

Fig 10 Functioning of MobilenetV2

Figure 11 shows the overall architecture of the MobilenetV2 approach to train and test the model for anomaly detection. After training and testing with MobilenetV2 is performed on the dataset, the model is trained. Trained model is then used to extract frames from test videos. Frames display the output whether the anomaly occurs or not in a given model.

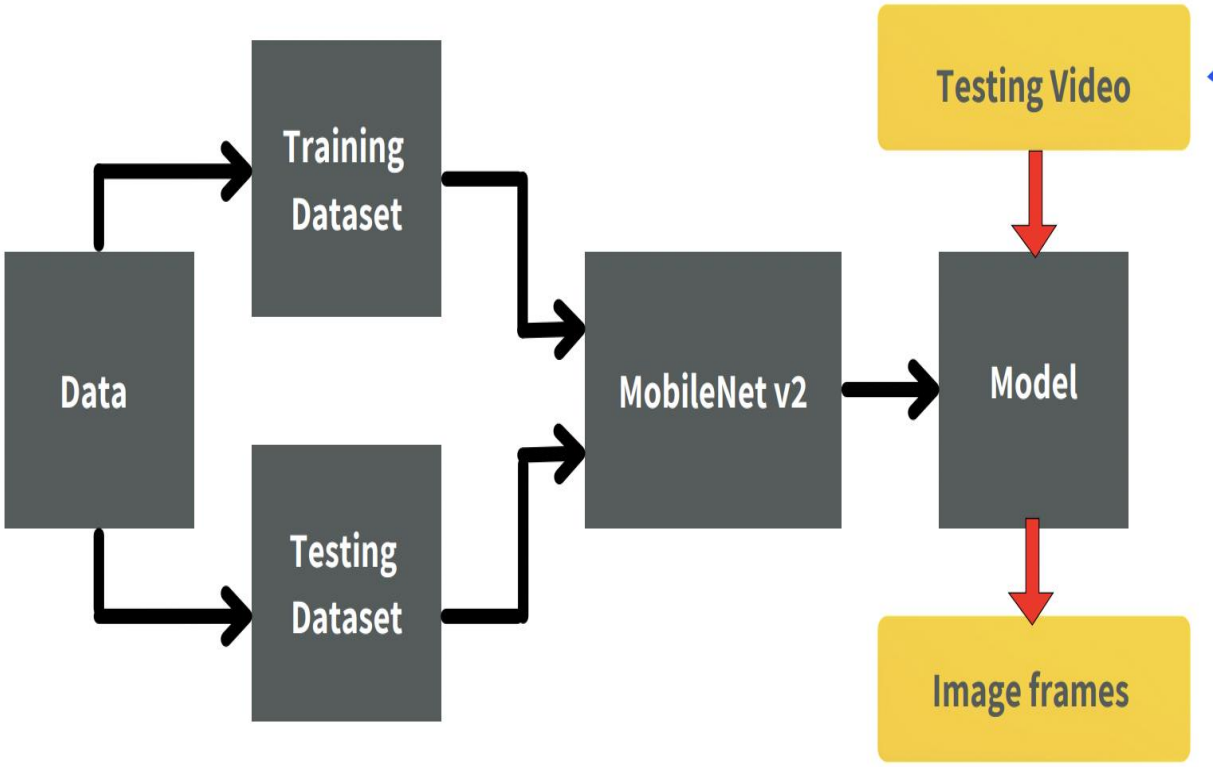

Fig 11 Architecture of MobilenetV2

Chapter 4: Implementation and Code Analysis

4.1 Implementation and GUI

4.1.1 MIL Approach:

All modules are integrated and compiled as a single entity. It is achieved by creating a model trained on weights into a Json file extension and importing the trained model into the GUI where it is applied onto the input video. Firstly, the inbuild Graphic User module of PyQt5 is imported and a popup window asks for the video to be uploaded in which anomaly is to be detected. Upon uploading the file, the features are extracted from the video and the file is converted into text format by C3D tool. Pre trained model is then applied to this video and the user interface displays graph of the anomaly.

The graph detects the anomaly at a particular frame along with the anomaly score or probability of anomaly. If the video is normal no change is seen in the result graph and it remains linear along x-axis. The window has multiple features to change dimensions of graph. The curve can be changed to a linear or log function depending upon the requirement. The curve color, width, style and line style all can be modified as per requirement. Labels along x-axis and y-axis can be named and renamed. Final values can be exported and saved. The graph can be exported as a jpeg file for further processing or extraction of imperative information. The GUI window has multiple features. Exact frames can be zoomed in upon and viewed. Pan functionality is also available. The complete window can be resized, and all image editing tools are provided in the GUI. This comprehensive Graphical User Interface provides the necessary results. The complete implementation is explained in various figures depicted below for a comprehensive understanding.

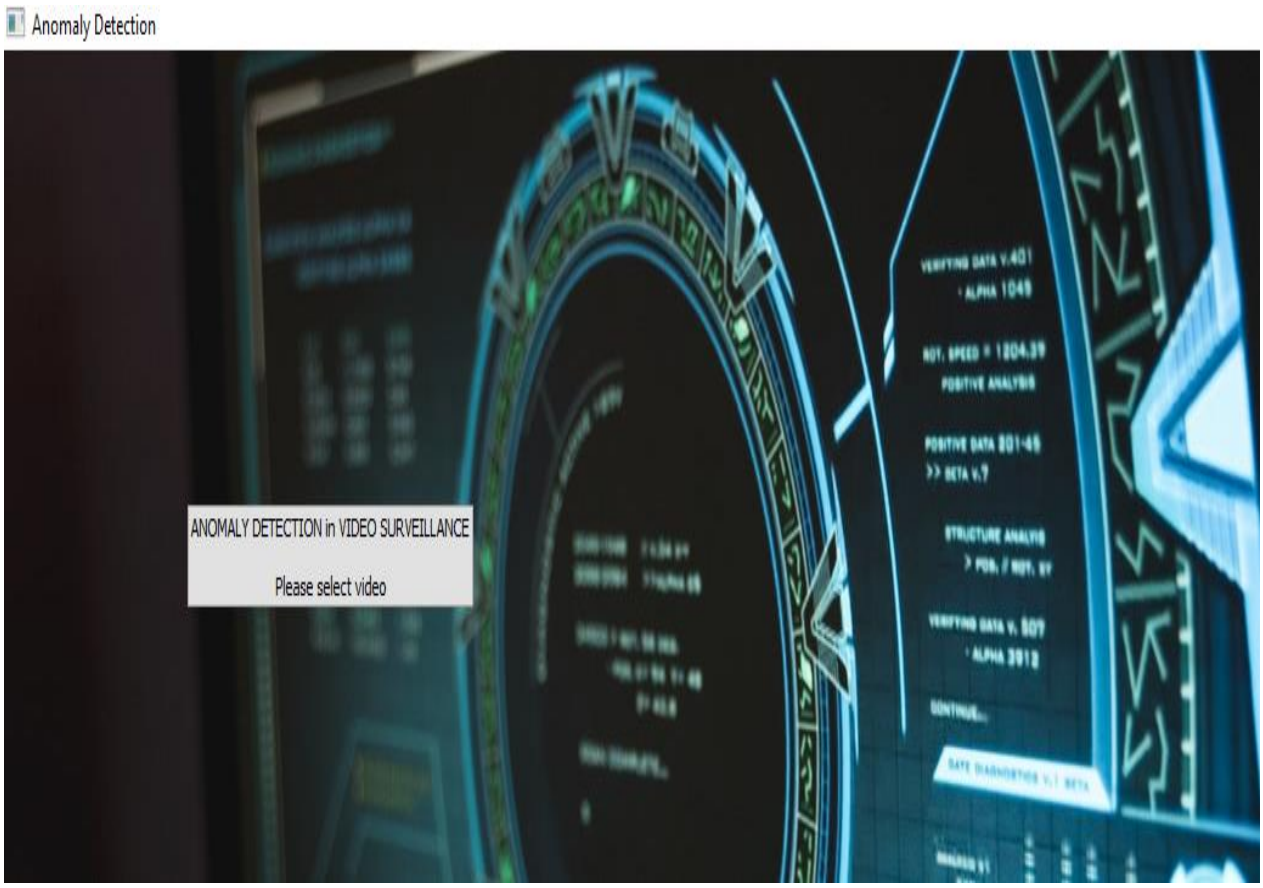

Fig 12 GUI Anomaly Detection System

Fig 13 Feature Extraction Tool

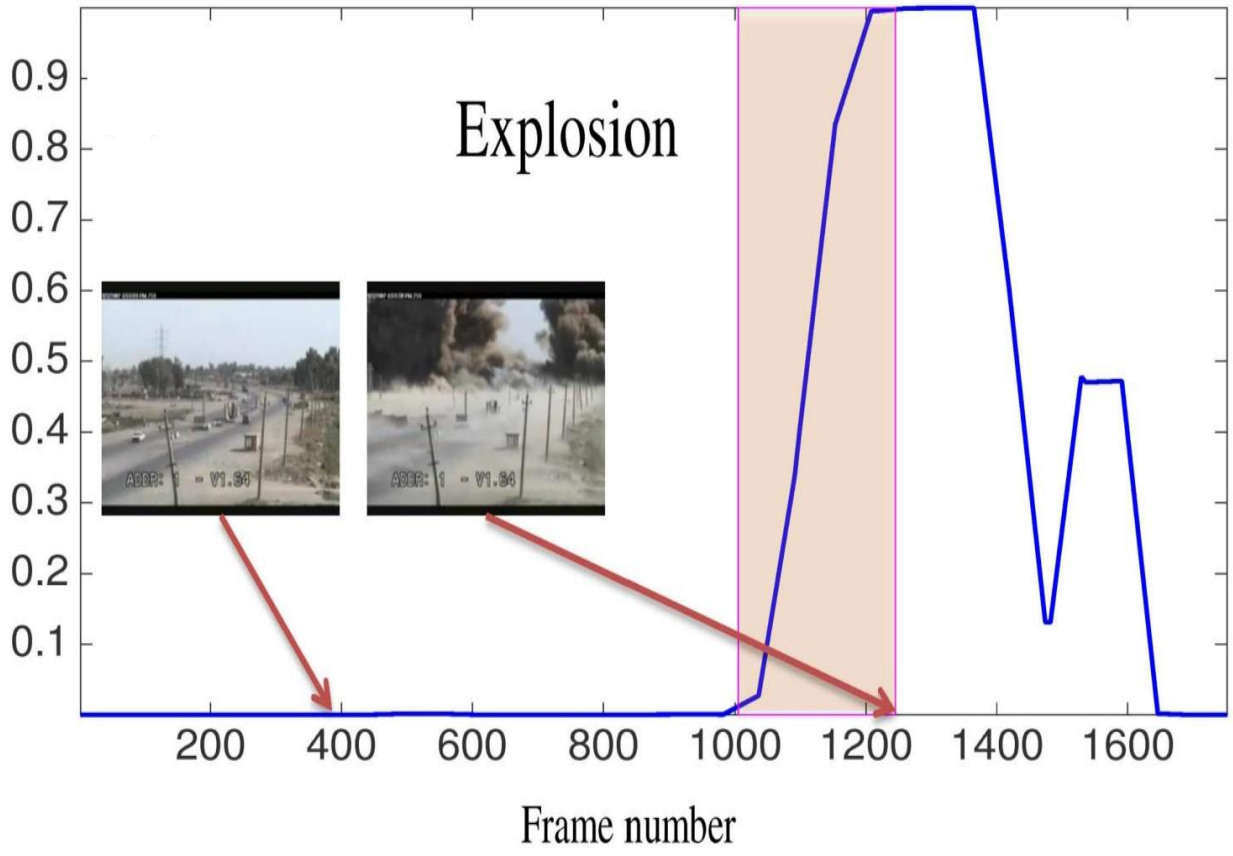

Fig 14 Graph as an output from ADVS

4.1.2 MIL Approach:

For the MobilenetV2 approach, we have implemented the training of the model on Google Colaboratory. Model is trained with 50 epochs on 350 violent and non-violent videos. After thorough training of the model on the real-time violence dataset, the model is created. After creation of the model, it's loaded on google colaboratory, testing video is uploaded from drive and frames are extracted depicting whether there's an anomalous behavior existing or not.

4.2 Code Analysis

The code has been implemented in python only using PyCharm software. Multiple libraries and dependencies as well as public domain software's have been imported into the environment and installed. The project is resource intensive and utilizes computation power and Graphical Processing Unit of the Machine it is being run on. Major imports are mentioned. TensorFlow and Keras have been extensively used. TensorFlow has been used to train model whilst utilizing its neural network capabilities to train and test our model. Keras library provides Python interface for our deep neural network. Keras acts as an interface for the TensorFlow library. Theano library has been used to calculate mathematical functions and apply multi-dimensional arrays efficiently. SciPy has been used for optimization and integration of mathematically complex functions. PyQt5 is the recent version of a GUI widgets toolkit. It has been utilized as it offers a Python interface in the form of a Graphical User Interface. It is by far one of the most powerful, and popular cross-platform GUI libraries. All widgets in the GUI have been offered by PyQt5. The project has various python extension files. TrainingAnomalyDetector.py, TestingAnomalyDetector.py and GUI.py all are separate files being integrated and compiled inside GUI.py once the main function is called.

4.3 Results

Results would be depicted in the form of graphs. The output is presented in the form of curve that depicts detection of anomaly. The beginning and termination of anomaly is also highlighted. The GUI has been explained elaborately in chapter 4.1 Implementation and GUI.

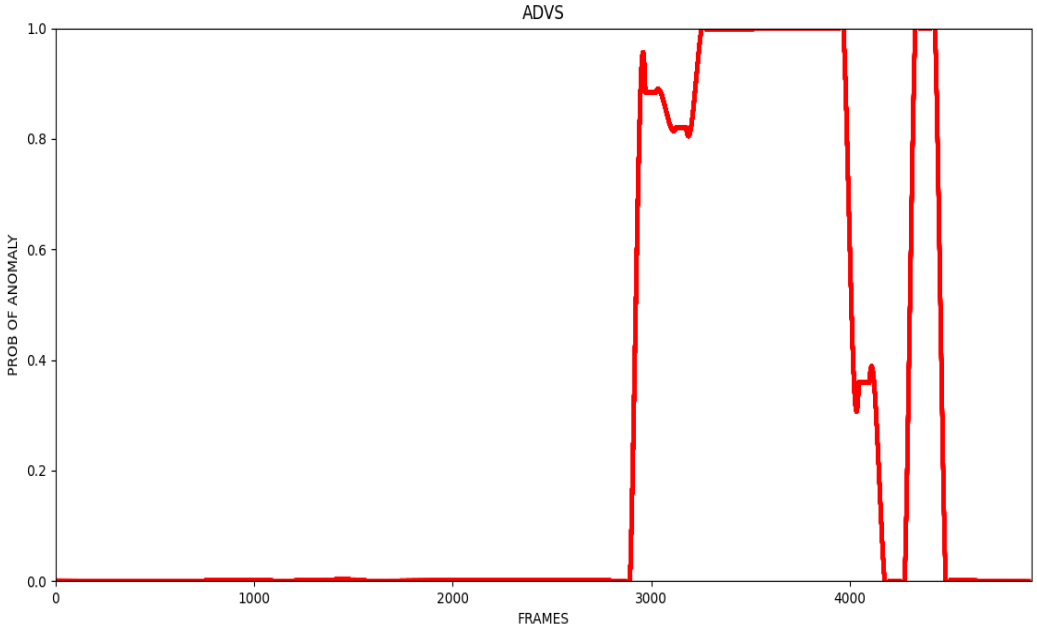

Fig 15 Anomaly Detection in Explosion Video

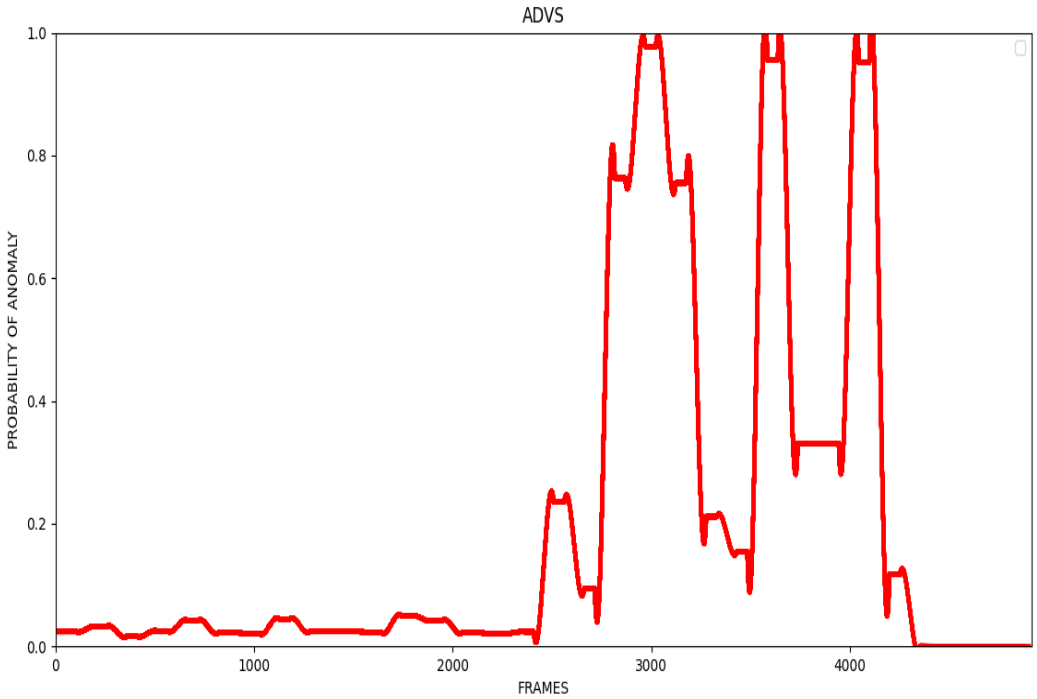

Fig 16 Anomaly Detection as Multiple Explosions

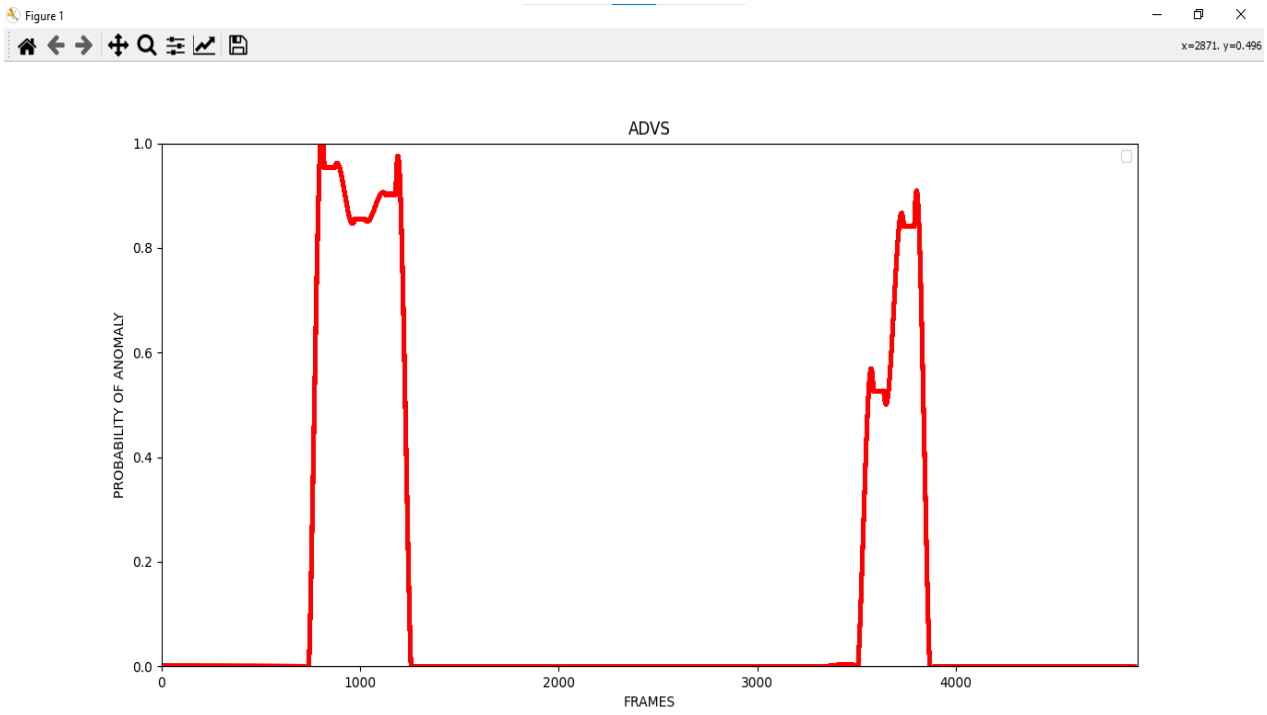

Fig 17 Anomaly Detection in Road Accidents

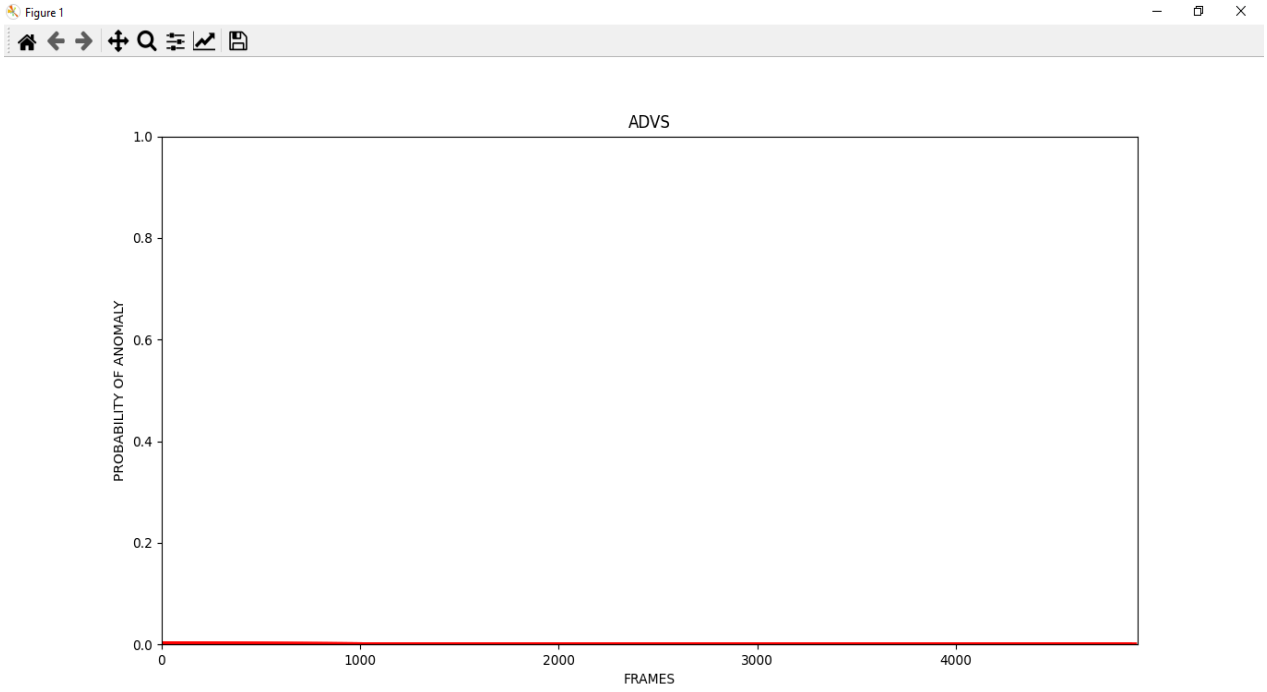

Fig 18 No Anomaly Detected in Normal Video

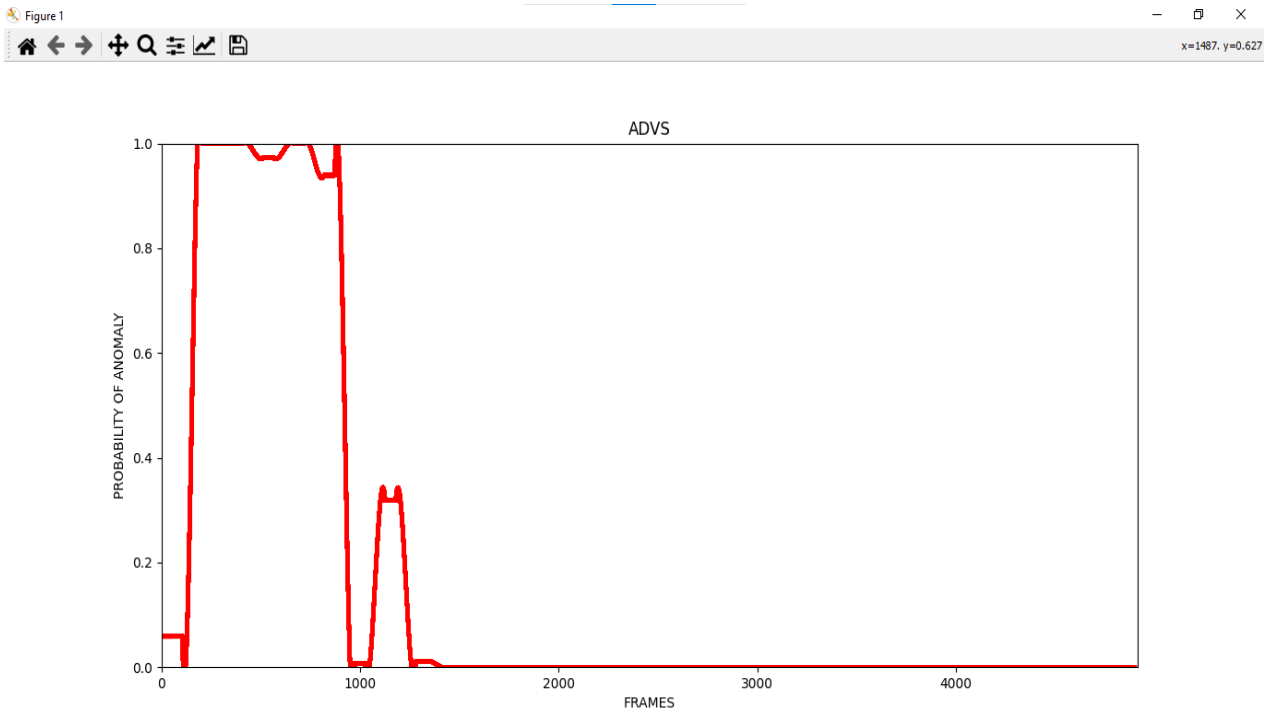

Fig 19 Anomaly Detection in Shooting Video

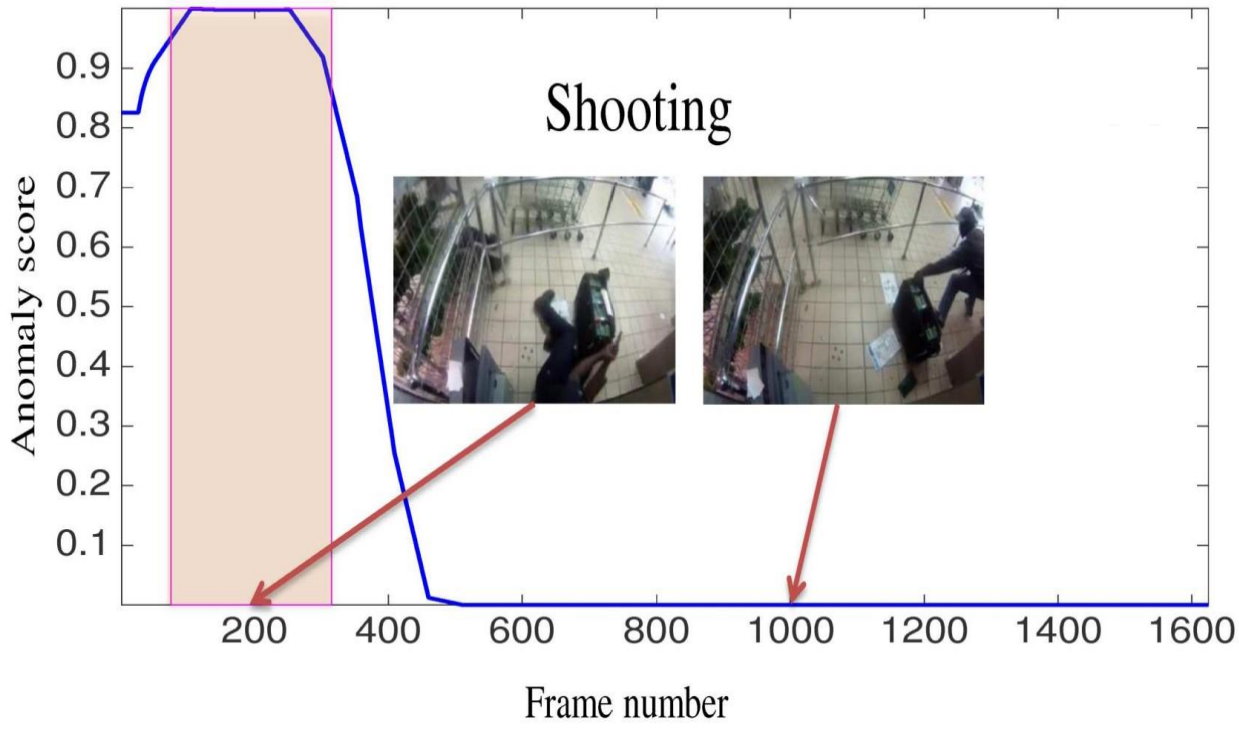

Fig 20 Anomaly Detection in Shooting Video (Labelled)

For MobilenetV2 approach, results are extracted in the form of video frames, displaying true or false for anomalous behavior. A video output file is also created with anomaly message depiction throughout the video. Figure 23 shows a random frame from an anomalous video showing violence as true.

Fig 21 Anomaly Detection (Violence) using MobilenetV2 Approach

Chapter 5: Future Work

In our final year project, we have designed, developed, and deployed Anomaly Detection system. The system works by uploading videos as an input. The system extracts feature from the video and converts those features in the form of a text file. This file is then fed to the Anomaly Detection trained model which performs computations on this text file and invokes the GPU that creates a graph. This graph displays a linear line along x-axis if the video is normal and has no anomaly in it. The graph paints a curve if any anomaly is detected. In the case of MobilenetV2 approach video frames are extracted and an output file is created which displays the presence of anomalous behavior. Although, the systems are trained and tested on anomalies mentioned above. Through rigorous training of model, the system is now capable of detecting all types of anomalies.

The video surveillance concept of the Anomaly Detection would be adopted in the future. This entails real time deployment of system onto camera processing machines or units. Where the camera feed would also depict graph along with feed. As soon as the anomaly would be detected the graph would raise along y-axis depending upon probability score. If the threshold of probability score is reached the alarm would be raised. Alerting operators in the vicinity or the manager through a notification. This future work is conceived to be a portable project in form of an application that could be deployed at any surveillance device.

Chapter 6: Conclusion

This project proposed a deep learning approach using artificial neural networks to detect real world anomalies in surveillance videos. Owing to the complex nature of these real-life anomalies. A successful attempt has been made to exploit both normal and anomalous videos. To avoid labor-intensive task of monitoring videos, the anomaly detection has been automated in the form of an Anomaly Detection System. We developed a general model of anomaly detection using TensorFlow and Keras libraries. To validate our approach, new massive scale anomaly dataset comprising of a various real-world anomaly is introduced. The results mentioned in chapter 2.2 depict our success on both approaches as our model performed significantly better than any other baseline model.

Anomaly Detection is considered to be one of the most challenging problems of machine learning. Anomaly detection on images is considered to be a challenge but on videos the challenge is multiplied by a factor. Through all difficulties we were able to develop our Anomaly Detection in Video Surveillance system.

References and Work Cited

1. http://www.multitel.be/image/research development/research-projects/boss.php.
2. Unusual crowd activity dataset of university of minnesota. In http://mha.cs.umn.edu/movies/crowdactivity-all.avi.
3. A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz. Robust real-time unusual event detection using multiple fixed location monitors. TPAMI, 2008.
4. S. Andrews, I. Tsochantaridis, and T. Hofmann. Support vector machines for multiple-instance learning. In NIPS, pages 577–584, Cambridge, MA, USA, 2002. MIT Press.
5. B. Anti and B. Ommer. Video parsing for abnormality detection. In ICCV, 2011.
6. R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic. NetVLAD: CNN architecture for weakly supervised place recognition. In CVPR, 2016.
7. A. Basharat, A. Gritai, and M. Shah. Learning object motion patterns for anomaly detection and improved object detection. In CVPR, 2008.
8. C. Bergeron, J. Zaretzki, C. Breneman, and K. P. Bennett. Multiple instance ranking. In ICML, 2008.
9. V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. ACM Comput. Surv., 2009.
10. X. Cui, Q. Liu, M. Gao, and D. N. Metaxas. Abnormal detection using interaction energy potentials. In CVPR, 2011.

11. A. Datta, M. Shah, and N. Da Vitoria Lobo. Person-onperson violence detection in video data. In ICPR, 2002.

12. T. G. Dietterich, R. H. Lathrop, and T. Lozano-Perez. Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence, 89(1):31–71, 1997.

13. S. Ding, L. Lin, G. Wang, and H. Chao. Deep feature learning with relative distance comparison for person re-identification. Pattern Recognition, 48(10):2993–3003,2015.

14. J. Duchi, E. Hazan, and Y. Singer. Adaptive sub gradient methods for online learning and stochastic optimization. J.Mach. Learn. Res., 2011.

15. Y. Gao, H. Liu, X. Sun, C. Wang, and Y. Liu. Violence detection using oriented violent flows. Image and Vision Computing, 2016.

16. A. Gordo, J. Almazan, J. Revaud, and D. Larlus. Deep image retrieval: Learning global representations for image search. In ECCV, 2016.

17. M. Gygli, Y. Song, and L. Cao. Video2gif: Automatic generation of animated gifs from video. In CVPR, June 2016.

18. M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury and L. S. Davis. Learning temporal regularity in video sequences. In CVPR, June 2016.

19. G. E. Hinton. Rectified linear units improve restricted boltzmann machines vinod nair. In ICML, 2010

20. T. Hospedales, S. Gong, and T. Xiang. A markov clustering topic model for mining behavior in video. In ICCV, 2009.

21. R. Hou, C. Chen, and M. Shah. Tube convolutional neural network (t-cnn) for action detection in videos. In ICCV,2017.

22. T. Joachims. Optimizing search engines using clickthrough data. In ACM SIGKDD, 2002.

23. S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi. Traffic monitoring and accident detection at intersections. IEEE Transactions on Intelligent Transportation Systems, 1(2):108–118, 2000.

24. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. In CVPR, 2014.

25. J. Kooij, M. Liem, J. Krijnders, T. Andringa, and D. Gavrila. Multi-modal human aggression detection. Computer Vision and Image Understanding, 2016.

26. L. Kratz and K. Nishino. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In CVPR, 2009.

27. Chandran junior Violence Detection in Real Time System June 2019

28. W. Li, V. Mahadevan, and N. Vasconcelos. Anomaly detection and localization in crowded scenes. TPAMI, 2014.

29. C. Lu, J. Shi, and J. Jia. Abnormal event detection at 150 fps in matlab. In ICCV, 2013.

30. R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In CVPR, 2009.

31. S. Mohammadi, A. Perina, H. Kiani, and M. Vittorio. Angry crowds: Detecting violent events in videos. In ECCV, 2016.

32. H. Rabiee, J. Haddadnia, H. Mousavi, M. Kalantarzadeh, M. Nabi, and V. Murino. Novel dataset for fine-grained abnormal behavior understanding in crowd. In 2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2016.

33. I. Saleemi, K. Shafique, and M. Shah. Probabilistic modeling of scene dynamics for applications in visual surveillance. TPAMI, 31(8):1472–1485, 2009.

34. A. Sankaranarayanan, S. Alavi and R. Chellappa. Triplet similarity embedding for face verification. arXiv preprint arXiv:1602.03418, 2016.

35. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res, 2014.

36. W. Sultani and J. Y. Choi. Abnormal traffic detection using intelligent driver model. In ICPR, 2010.

37. Theano Development Team. Theano: A Python framework or fast computation of mathematical expressions. arXiv preprint arXiv:1605.02688, 2016.

38. D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri. Learning spatiotemporal features with 3d convolutional networks. In ICCV, 2015.

39. Real-world Anomaly Detection in Surveillance Videos – Center for Research in Computer Vision (ucf.edu)

40. J. Wang, Y. Song, T. Leung, C. Rosenberg, J. Wang, J. Philbin, B. Chen, and Y. Wu. Learning fine-grained image similarity with deep ranking. In CVPR, 2014.

41. S. Wu, B. E. Moore, and M. Shah. Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In CVPR, 2010.

42. D. Xu, E. Ricci, Y. Yan, J. Song, and N. Sebe. Learning deep representations of appearance and motion for anomalous event detection. In BMVC, 2015.

43. T. Yao, T. Mei, and Y. Rui. Highlight detection with pairwise deep ranking for first-person video summarization. In CVPR, June 2016.

44. Waqas Sultani, Chen Chen, Mubarak Shah, Real-world Anomaly Detection in Surveillance Videos, Cornell University Library, arXiv:1801.04264 [cs.CV], [v1] Fri, 12 Jan 2018.

45. https://www.kaggle.com/datasets/mohamedmustafa/real-life-violence-situations-dataset

46. https://towardsdatascience.com/review-mobilenetv2-light-weight-model-image-classification-8febb490e61c