Anomaly Detection in Video Surveillance using Deep Learning

Harshitha. G[1], Hemanth Kumar. A. R[2]
M.Tech Scholar[1], Professor & Head[2]
Department of Electronics and Communication Engineering
Bangalore Institute of Technology, Bangalore-560004, India

Abstract:

Proficient anomaly detection in surveillance videos across different conditions addresses a significant challenge in PC vision. This work proposes a background deduction approach using deep learning techniques of residual neural networks equipped for recognizing from various action objects to various sizes by foreground segment per pixel. The proposed algorithm accepts input as a reference source and as a target frame, needs to be adjusted temporarily and generates division instructions of the same spatial target. Tests display serious execution in the examined dataset, as well as ongoing ability. In recent years, surveillance cameras have been introduced in various fields. Investigating the data obtained with these cameras can produce fascinating pieces in occasion forecast, checking based on web and objective driven investigation applications including irregularities and interruption identification. These days, different Artificial Intelligence methods have been utilized to recognize peculiarities, among them are convolutional neural networks utilizing profound learning procedures further developed the identification precision altogether. The objective of this project is to propose another similar technique in view of deep learning methods for identifying the threatening actions from the video surveillance cameras.

Keywords- Deep Learning, Multiple Instance Learning (MIL) Algorithm, Video, Detection, Anomalies, Alert.

I. INTRODUCTION

Computerization of Video Surveillance is acquiring broad interest as of late, taking into account the security issues of the public. Efficient and accurate event capture has become an important issue in modern computer imaging technology. The development of contemporary deep learning techniques has significantly raised interest in this area and enhanced object capture accuracy. The question of whether to detect anomalies using supervised or unsupervised machine learning/deep learning algorithms is still up for discussion. An anomaly is an occurrence of an event or action that is out of the ordinary, unexpected, unpredictable, odd, and hence dissimilar from the regular pattern [1]. Detecting anomalous scenes during normal data playback can be an important and unique application. Additionally, the environment and conditions around the anomaly have a total impact on the anomaly identification procedure. The portrait video format has an impact on modelling and rendering. Events in the real world are complex and difficult to understand. Anomaly-based detection algorithms have achieved the certain level of accuracy under specified conditions, however the technique is dependent on both internal and external elements, such as object illumination, motion direction, trip velocity, gesture, and closure [2].

Currently, Closed Circuit Television (CCTV) camera video is being generated faster every minute online with more cameras installed in public areas to improve efficiency, security, and safety following crimes and terrorist acts. According to reports, these CCTV cameras often keep an eye on places like malls, hotels, roadways, banks, and governmental structures [3,1]. However, it is reliable to check hundreds of employee-operated security cameras for strange activities. A potential answer to this issue is the creation of clever computer algorithms that can automatically identify unusual occurrences in the video environment. The identification of the anomaly in this research is based on the footage from security cameras. It should be highlighted that because video involves detecting methods and requires further video processing [4, 1].

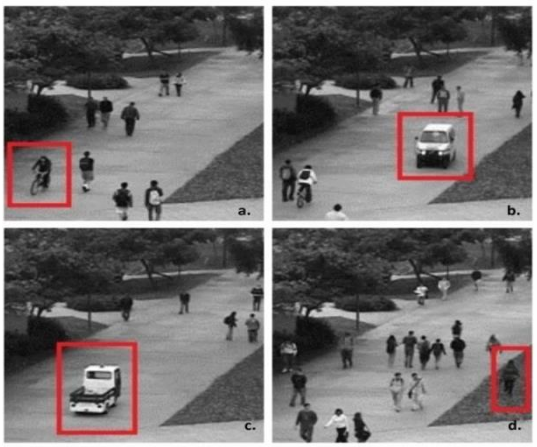

Figure 1: Illustration of Anomalies Detected in video supervision scenes

Figure 1 illustrates the abnormal behaviors example in a video from surveillance camera scenes, which includes access to a limited area, leaving unfamiliar packages, improper movements of vehicles, people and in misguided direction, which is seized by the camera surveillance systems at airports, vehicle park zones, stations and other public places in general [5]. The processing of information from a surveillance- cameras in full footage is a big challenge and also difficult. Assuming this interaction takes place on the web, the complexity increases. Using cutting-edge machine learning techniques like deep learning is perhaps one of the finest ways to manage and interpret this data. End-to-end systems are what make these kinds of cycles, which frequently include massive data, beneficial. The end-to-end system automatically carries out feature extraction.

The main contribution of this study is the use of deep learning algorithms in all facets of anomaly identification. In Section II related works of the project is discussed, Subsequently, in Section III the suggested new approach is thoroughly discussed through methodology. Further, in Section IV results and its discussion will be explained in brief and at last is in Section V, conclusions are drawn and the future scope is presented.

II. RELATED WORKS

This section briefly discusses on the conception of Machine Learning/ Deep Learning (ML/DL) techniques and further about output approaches along with datasets, which are useful in detecting the anomalies and to carry out the overall proposed system.

The kind of input data significantly affects the deeper neural network structure that is used for deep-anomalies detection techniques. Labels show whether a certain data occurrence is an outlier or not. Since anomalies are uncommon, it might be challenging to assign labels to them. Anomaly behaviour can also alter over time. For both typical and atypical data scenarios, supervised DAD (Deep-Anomaly Detection) comprises thorough training in supervised binary categorization or multiple categories utilizing labels. Supervised learning is the most widely used kind of machine learning and deep learning. A labelled database is used in supervised learning to train algorithms that predict or rank outcomes [6, 1].

A. *Stages of ML/DL*

There are five stages involved in the software part of the proposed system, they are: Data collection, Pre-processing data, Data splitting (Train & Test), Model building and Model evaluation/deployment [7]. First stage, finding methods and resources to gather accurate, complete data, interpreting it, and utilising statistical methods to analyse the findings constitutes the collection of data. One of the key steps in the machine learning or deep learning process is gathering data to train a model. The accuracy of the predictions provided by ML/DL systems is only as good as the training set of data. Here, the data collection tools consists of two phases such as Quantitative and Qualitative. Quantitative tools include online web, live, telephone, Fedex mail and central location interception. Whereas, qualitative tools include online forums, communities, web surveys, groups- triads- dyads and in-depth interviews (IDI) [8].

Second stage, pre-major processing's goal is to transform raw data into a format that is acceptable for ML/DL and further minimizing the data. Figure 2 shows the Minimizing Data Leakage.

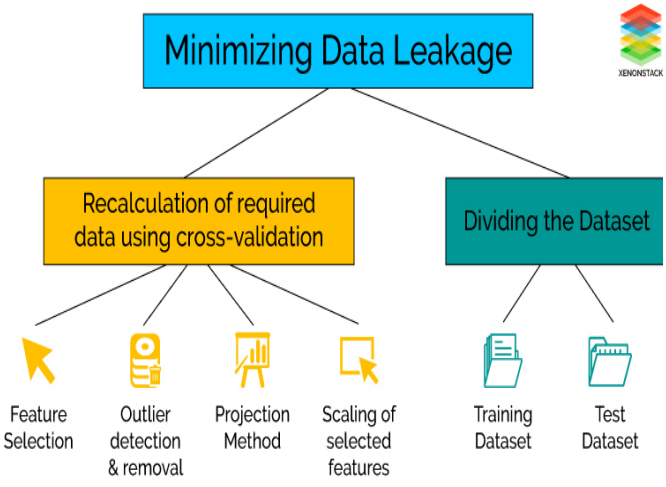

Figure 2: Minimizing Data Leakage

Obtained results from the applied ML/DL model are more accurate when the data is clean and organized. Raw real-world data and photos are frequently unreliable, lacking in certain behaviours or patterns, and prone to a number of mistakes. As a result, after being gathered, the data is preprocessed into a format that the machine learning algorithm can utilize for the model. In the world of data protection, the term data leaking relates to the unauthorized movement of information outside of a protected facility such as a data center. However, given the importance of keeping details concerning the forecast completely separated from the retraining and model development stages, this protection approach is actually relatively appropriate for our machine educational environment. One essential guideline is to make sure that any data pre-processing is done separately for each cross-validation sweep [9, 8].

Third stage, Splitting datasets (first training split and then testing split), is a technique for testing the effectiveness of machine learning systems. It may be used again for classification or regression issues as well as any supervised learning technique, whenever a model is being used to predict data that wasn't utilised to train it, data set splitting processes are used to assess how well ML or DL algorithms perform. Objective evaluation of prediction performance requires splitting of dataset [10].

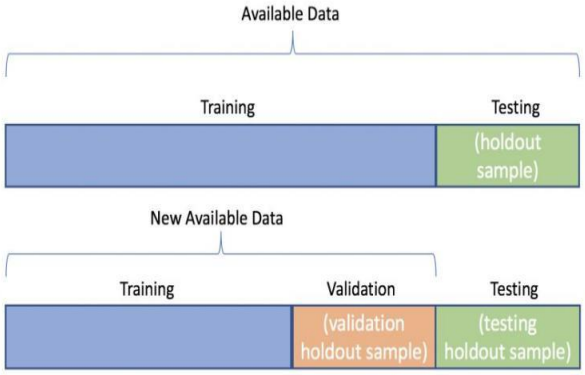

Figure 3: Dataset Splitting.

As shown in figure 3, the process involves splitting a dataset into two subgroups. The first set, often referred as the train the data set, is utilized to fit the model; the second sub - set is not being used to train the model. Alternatively, predictions are made and matched to the projected values once the model has been provided the input from the data set. The second set of data is referred to as the test dataset. Usually, both train data set and test data set have a ratio rate of 80percent in terms to 20%, respectively. Following that, a re-training set is divided, with 20% of it being utilized to create the valid set. Every-time one can train a ML/DL model, but it is unable to do so on a single dataset, even then we are unable to asses the models performance. For this reason, it is necessary to split our source data into training, testing, and validation datasets [10].

Fourth stage is building a model, in this step, the user trains multiple models to determine which of them provides the nearly accurate predictions. Basically, an ML/DL model is detailed as a mathematical representation of the results of the training process. The study of different model algorithms that mechanically become better with practise, stale data, and model development is known as ML/DL. First of all, these models are trained on datasets, then algorithm fed to make inferences about data, next extract the model from the stream data, and learn from that data. After these models are trained,

they can be used to predict the invisible data set [7].

Fifth stage is about ML/DL model evaluation or deployment is the procedure of guiding a completed training model in a live environment where it can be used for its designed purpose. In the below Figure 4, one can notice that the following should be taken into account when the model is trained. This is This is necessary to select a model from many trained models.

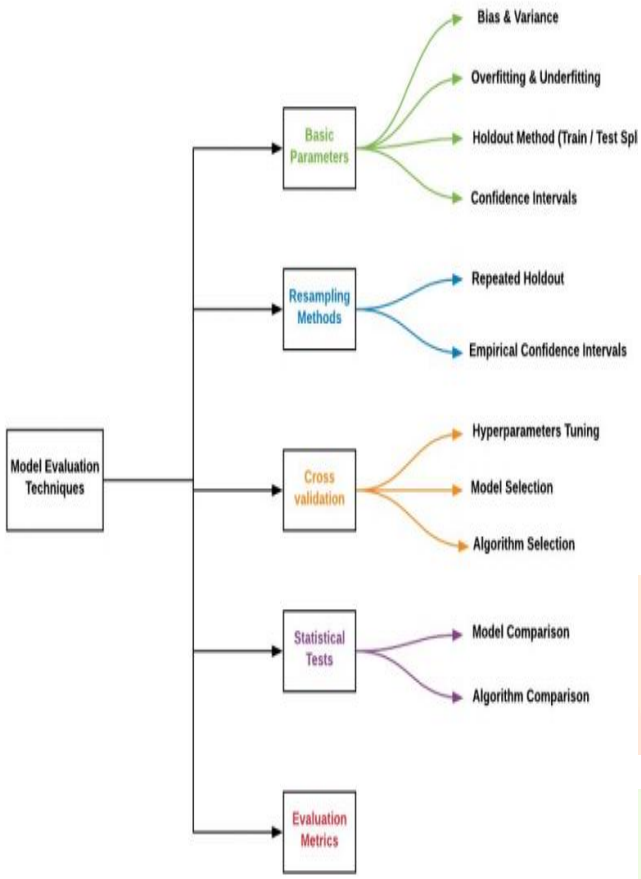

Figure 4: Different types of Model Evaluation

Models can be deployed in a variety of environments and are often integrated into applications through an API that is accessible to end users. Model evaluation plays a crucial role in developing any ML/DL model. Building a predictive. In this, ML/DL model without checking for errors, cannot be counted as a fit model. To enhance the accuracy, we need to check on the metrics and make improvements accordingly until we get the desired accuracy rate. To start using a model for actual decision making, it needs to be effectively deployed in production [7].

After the completion of the five stages of machine learning/deep learning techniques, it is important to find out how abnormalities are recognized. This is a crucial component of anomaly detecting technologies. The outcomes generated by anomaly detection techniques often come in the form of a binary label or an anomaly score [1].

There are numerous public databases available that verifies surveillance and anomaly detection algorithms. UCSD, CUHK, Avenue, UMN, and Subway are most widely used datasets to detect anomalies. In this proposed project, the UCSD dataset was used more than the others. Events from various crowded settings, ranging from few to packed, are included in the UCSD database. The dataset depicts a variety of situations, including strolling on a road, on grass, operating a car on a sidewalk, and surprising behaviours like rollerblading. The dataset used in this proposed project is from UCSD [1].

III. PROPOSED METHODOLOGY

The method proposed in this project uses deep learning algorithms to find anomalies and its activity patterns in surveillance footage. The proposed block diagram for anomaly detection in a deep learning-based video surveillance system is shown in Figure 5. A Microcontroller, GSM, LCD, Buzzer, and a video processing unit make up the framework. The system's controller is connected to GSM, LCD, and buzzer communication devices, which functions as its brain. The Buzzer is placed in the control room. This controller is also connected to a video processing unit system which is located in a guarded areas like any public area or in any military facility guard area.

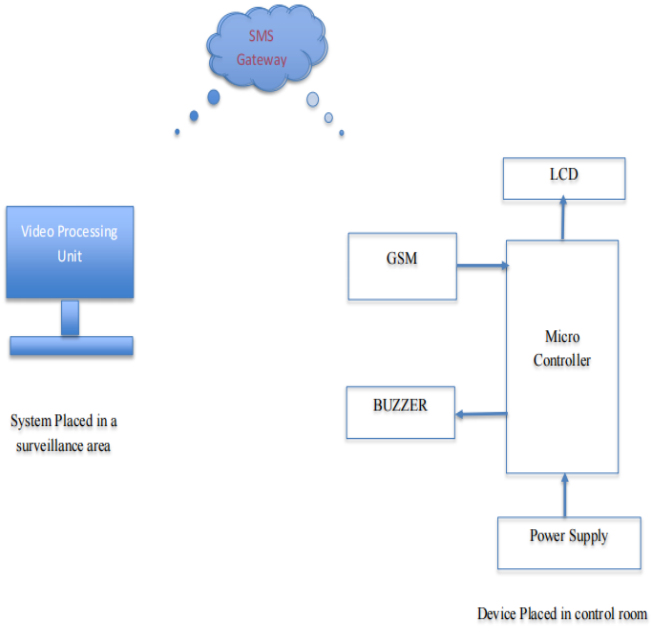

Figure 5: Functional block diagram of anomaly detection in video surveillance

A. Anomaly Detection in Video Processing Unit

For the purpose of identifying anomalous occurrences in video footage and categorizing them as either normal or anomalous, discrete machine learning or deep learning technologies are utilized. Figure 6, depicts the anomaly detection in broad strokes. Visual sensors in the monitoring region gather the data during this procedure.

Following that, this raw visualization data will go through processes like feature extraction and pre-processing. The generated data is sent into a modelling system, which analyses the behaviour to see whether it is abnormal and tracks the learning process that was used to predict the intended behaviour. The Multi-Instance Learning (MIL) Algorithm is the learning technique employed in this project. On the basis of labeled bags used as training data, this kind of algorithm categorizes occurrences or unseen bags [4].

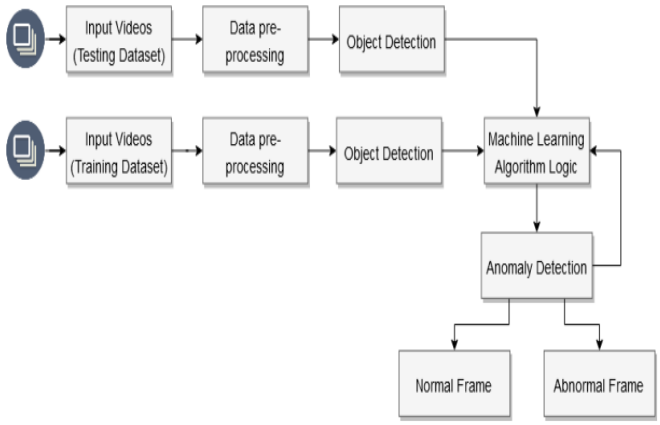

Figure 6: General overview of the Anomaly Detection

Once the anomaly has been detected from the processing unit, next step is to calculate the accuracy rate of the model's algorithm and this can be done using the confusion matrix.

Basically, confusion matrix is a method for measuring a classification algorithm's performance. If there are more than two classes in our dataset or if there aren't an equal number of observations for each class, the classification accuracy alone may be deceptive. We can gain a better understanding of the categorization model's successes and failures by calculating a confusion matrix. The matrix has two dimensions: actual values and predicted class values, as well as the number of total of predictions, as shown in the figure 7. Actual values are the real values for the provided data, whereas predicted values are the values that the model predicts.

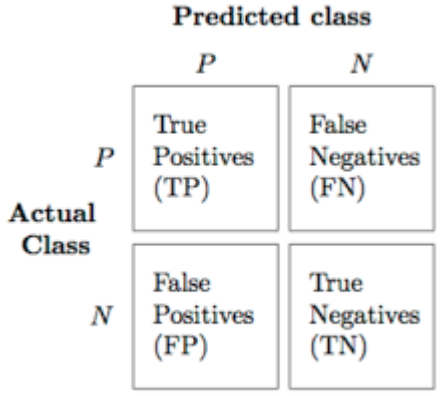

Figure 7: Vectored implementation of Confusion Matrix

Using this confusion matrix, we then calculate the model's accuracy as well as other things. The following computations are provided:

1. Classification Accuracy: It is one of the crucial factors in figuring out how accurate a classification problem is. It specifies how frequently the model predicts the right result. The number of accurate predictions made by the classifier divided by the total number of predictions made by a classifier can be used to compute it. The following is the formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

2. Misclassification rate: It also goes by the name "error rate" and describes how frequently the model makes incorrect predictions. The ratio of wrong guesses to all of the classifier's predictions can be used to compute error rate. The following is the formula:

$$\text{Error rate} = \frac{FP+FN}{TP+FP+FN+TN}$$

3. Precision: It can be interpreted as the number of accurate outputs produced by the model or as the proportion of correctly anticipated positive classes that actually occurred. Using the formula below, it can be calculated:

$$\text{Precision} = \frac{TP}{TP+FP}$$

4. Recall: It is referred to as the percentage of the total positive classes that our model accurately predicted. There must be a significant recall.

$$\text{Recall} = \frac{TP}{TP+FN}$$

5. F-measure: It is challenging to compare two models that have low precision but good recall, or vice versa. F-score can therefore be used for this purpose. This score enables us to simultaneously assess recall and precision. If the recall and precision are equal, the F-score is at its highest. Using the formula below, it can be calculated:

$$\text{F-measure} = \frac{2*Recall*Precision}{Recall+Precision}$$

Confusion Matrix also uses any one of 2 key terms such as; Null error rate or a ROC curve. In this project work, we calculate the ROC curve mainly. The performance of a classifier for each potential threshold is shown on a graph called the Receiver Operating Characteristic (ROC). The real positive rate (on the Y-axis) and also the false positive rate are shown on a graph (on the x-axis), as shown in the figure. 13 in the Section IV.

B. Alert through SMS Gateway

The proposed device will be placed in the control station or room. As soon as the power supply is connected to the system micro-controller, LCD and GSM/GPRS will get initialized. When a video processing unit system detects an abnormality, an gateway coupled with the position of the camera, at the same time siren will be turned on in the control room indicating the problem.

C. Flow Diagram for Detecting Video Anomalies

This approach consists of two primary parts. Anomaly detection makes up the second half of the component, whereas feature extracting and learning makes up its initial half. Despite these two elements, estimating and reducing background noise requires a pre-processing step. This approach also includes two primary parts, like other machine learning techniques: a training phase and a testing phase, as seen in Figure 8 below.

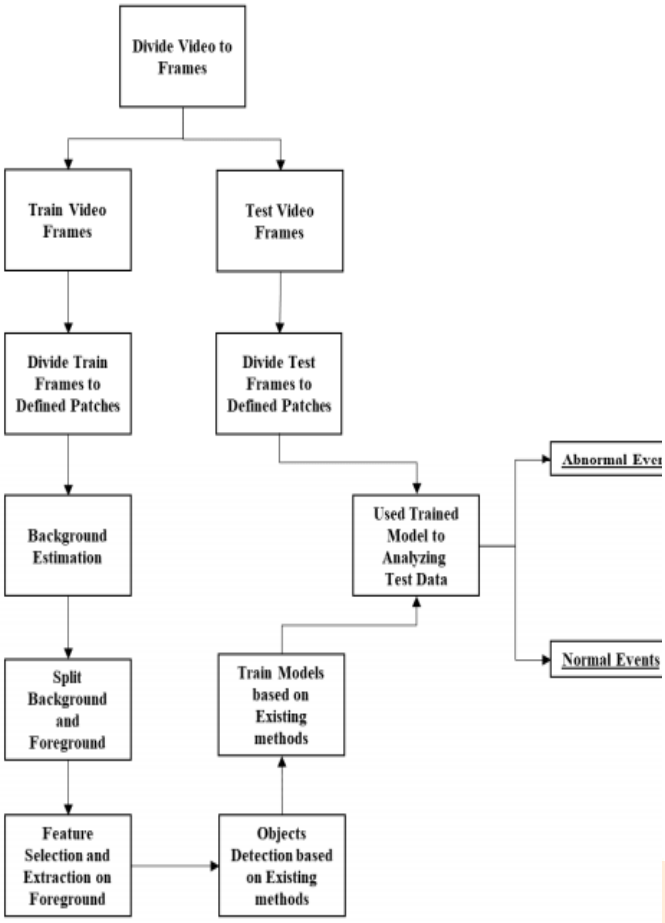

Figure 8: Video Anomaly Detection Flow Diagram

The features are trained using portions of the database that only include normal/regular frameworks in the train phase, as well as from a model utilised in other portions of the database that contain unusual/unusual frames from the test phase. Figure 8 illustrates the four basic categories that learning qualities fall under. Some feature extraction procedures are carried out on individual frameworks, while others that are based on patch frameworks will be altered to reduce the time and expense associated with training. The first component is the event associated with the acquisition of each item frame from a video; the achieved result (detection point) is formed by comparing each frame with the prior frame and the next frame.

Finally, a final score is determined based on each comparison of pictures at a reasonable pace. Size in relation towards the volume of stuff in each frame makes up the second criterion. The third component is the flow-based movement of objects within the patch frames, which results in an optical flow and video
sequence and another anomaly score.

D. Case Study

Before discussing about the case studies, there are 3 important parameters of this study to be considered, they are: Image frame, Anomaly score and Segment.

1) Image frame: The input video is divided into frames and these frames are considered as images, in this particular project work. For the case study, particular anomalies detecting frames has been taken for image. Here, six types of images are considered; image 1 is road accident, image 2 is assault, image 3 is shooting, image 4 is fighting, image 5 is explosion and the last image 6 is the normal image like walking or running.

One minute video input is taken here, in which per second has 30 frames are being extracted. Therefore,1 minute is 30 frames * 60 seconds which gives 1, 800 frames.

2) Anomaly score: The probability calculation alone determines the anomaly score. To make the anomaly more understandable, the descriptions and even the average value are condensed pieces of contextual information. The anomaly score, which ranges from 0 to 100 and in this study is represented in decimal points between 0.20 and 0.55, represents the significance of the abnormality relative to other anomalies that have been observed in the past. Low scoring values are displayed with the extremely abnormal results. A critical interval that needs more examination is one with a higher anomaly score.

3) Segment: In Image processing, Image Segment or segmentation is a technique that divides a digital image into various subgroups known as Image segments, which serves to simplify future analysis or processing of the image by decreasing the complexities of the original image. Basically, segmentation is the process of giving pixels labels. In this work, one segment has 16 frames. There are three characteristics of the Segments —spectral detail, spatial description, and minimum segment size— have a significant impact on the features of the picture segments. The level of information that defines an interesting characteristic can be changed.

To train and test a model, the datasets used in this project includes accident, explosion, shooting, assault, fighting and a normal data. The following are brief case study on these datasets.

i. Road accident

The first case is about detection of common road accidents in public places. A road accident is that any injury brought on by collisions that start, end, or involve a vehicle that is partially or entirely on a public road. Instead of being mechanical, human faults plays a major role in the rising number of road traffic accidents. The main causes of traffic accidents are drunk driving, excessive speeding, disregard for traffic laws, and reckless driving.

Figure 9(a): Image 1 of Road traffic accident

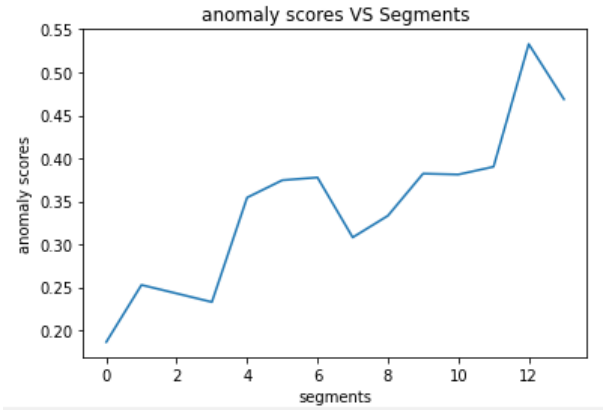

Figure 9(b) Accident detection.

The figure 9(a) shows the road traffic accident and 9(b) is the obtained graph when the accident detected. The graph is anomaly scores vs Segments where x-axis is segments of the image frames and y-axis is the anomaly scores. In every segment, whenever a anomaly detected, there is a spike or fluctuation in anomaly score as shown in the figure 9(b).

ii. Crime scenes

The second case is crime detecting. Here, crime scenes includes assault, fighting and shooting in the public or other places. The crime scene is a location where the crime has taken place along with its immediate surroundings. The act of committing bodily injury or unwanted physical harm to another person, or, in some legal definitions, the threat or attempt to do so, constitutes an assault. While fighting is a violent dispute or fight between 2 or more people, shooting is the process or act of firing a bullet from a ranging weapon (such as with a gun).

Fig.10 illustrates the area where 3 various crimes occurs and fig.11 shows the graphs upon crime detection.

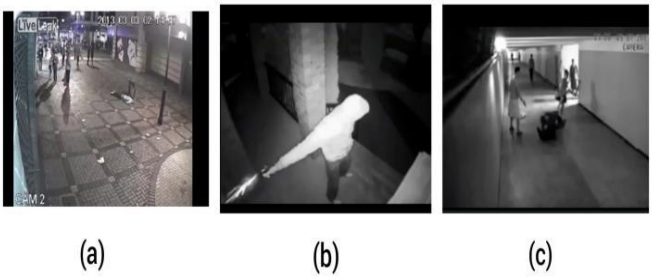

Figure 10: Crime scenes- (a) Image 2- Assault and (b) Image 3- Shooting and (c) Image 4- Fighting.

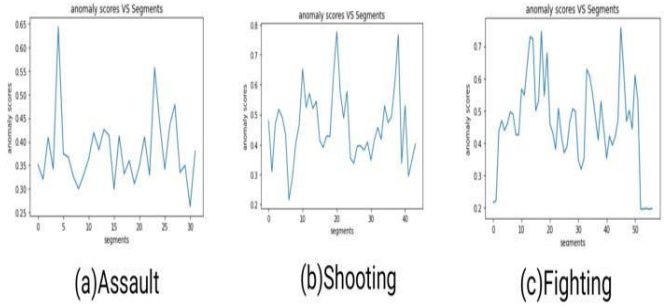

Figure 11: Crime scenes detection

The graph demonstrates anomaly scores vs segments, with segments created from a set of image frames on the x-axis and anomaly scores on the y-axis. Every time an anomaly is spotted in a segment, the anomaly score will spike or fluctuate, as seen in figure 11.

iii. Explosion

Third case discussion is on detection of explosion around public places. An explosion is a sudden increase in volume accompanied by a very strong energy release that usually results in the production of high temperatures and the emission of high-pressure gases.

Figure 12(a): Image 5- Explosion scene

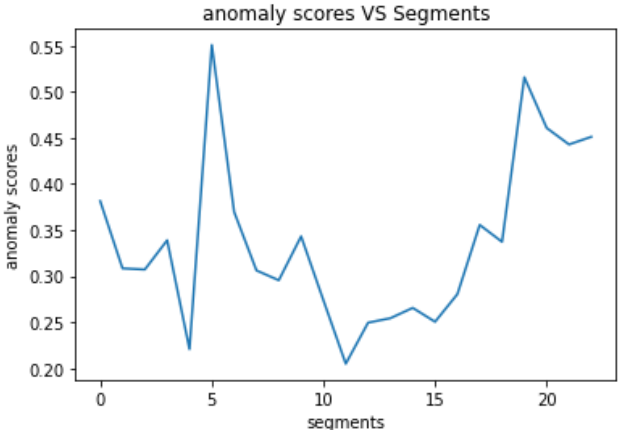

Figure 12(b): Explosion detection.

A explosion occurring on a public road is as shown in the figure 12(a) and in fig.12(b) illustrates the graph upon detecting anomaly that is explosion. Results from the applied machine learning model are more accurate when the data is clean and organized. On the graph, segments made from a collection of image frames are

plotted against anomaly scores in x-axis, and further that are plotted along the y-axis. This anomaly score will attain peak or fluctuate each time when an anomaly is detected in a segment, as seen in figure 12(b).

Another case is about the normal event where no anomalies are detected and normal event is nothing but a regular pattern that are usual, for example walking or running, which considered as the Image- 6.

IV. RESULTS AND DISCUSSION

The application which is enabled with the proposed system is discussed and analyzed throughout the results. The outputs of the CNN architecture-based MIL algorithm is carried out through Python version 3.7 platform, to detect the abnormal objects and events using the already trained datasets. By using GSM with SMS gateway feature, LCD with display feature and buzzer with siren feature, which all integrated with PIC micro-controller, this project provides an automatic anomaly detector for video surveillance. The proposed system is built as shown in the figure 5. When a input power supply is given to the micro-controller of normal system, then the GSM and LCD will get initialized. If any anomalies detected in processing unit, alert SMS with location of the nearby surveillance camera will be sent and siren turned on. Therefore, upon receiving the alerts, the controller will display "Alert!! Anomaly found" or else "Normal". This will alert the person incharge to take the required actions.

A. Training phase and Testing phase

The input video films are collected from UCSD databases and YouTube recordings of public sites. Around 350 videos of various abnormal and normal behaviors have been

gathered. Frames are taken from the recorded videos as part of the preprocessing procedure. Our system uses CNN with the MIL algorithm as a pre-trained model, and it uses what it has learned to address the issue at hand. It was used to train our datasets. For testing, CCTV video footage of various scenarios lasting about a minute is collected from frequent public spaces and processed into frames. The trained model receives the stored frames and categorizes the video into anomalous and normal behaviour.

B. *Quantitative Analysis*

A receiver operating characteristic, which is a graphical depiction of the overall representation of the entire trained datasets, is shown in Fig.13 below.

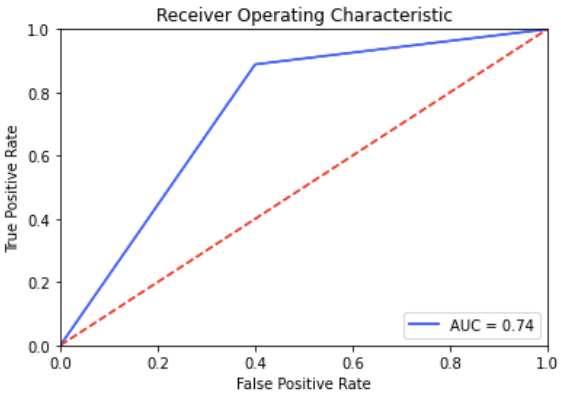

Figure 13: Receiver operating characteristic.

Around 300 varied normal videos, as well as about 50 abnormal videos, are used in this work. This graph(fig.13) demonstrates the percentage of successfully predicted and inaccurately predicted normal videos out of 300 regular videos. Here, a model has been built using 300 typical clips, and the percentage of accurately predicted videos, or the auc rate, is 0.74. The characteristic graph is used to calculate the average normal video rate rather than for algorithm performance. This characteristic analysis is carried out in order to determine the model's performance level in terms of the rate of true and false positives. The characteristic graph used to calculate the average normal video rate rather than for algorithm performance. This characteristic analysis is carried out in order to determine the model's performance level in terms of the rate of true and false positives.

$$\text{Accuracy} = \frac{\text{Correct predictions}}{\text{Total predictions}} * 100$$

Table 1 below, shows the quantitative analysis results obtained on the video images before and after using MIL on base algorithm. The images which were abnormal were detected and have been noted. The architecture of the suggested system has been significantly improved, according to the paper, as compared to previous pure software- based anomaly detectors. From the table 1, a image 1 is a accident class shown in fig.9, image 2 is the assault class, image 3 is fighting class and image 4 is the shooting class, is shown in the fig.10 with there respective graphs in fig.11, also image 5 is a explosion class as shown in fig.12, and image 6 is a normal class is a usual data (like walking or running); which all discussed briefly in the case study of Section III.

TABLE 1: Quantitative Analysis results.

Sl.No. (Video Images)	Class	Base Algorithm detection percentag-e	MIL detecti-on percent-age	Results of Analysis
Image 1	Road accident	43%	87%	Improved Accuracy and detection of vehicles
Image 2	Assault	51%	81%	Improved accuracy
Image 3	Shooting	48%	92%	Improved detection of pistols, Knife and weapons and Accuracy
Image 4	Fighting	51%	80%	Improved accuracy
Image 5	Explosion	46%	78%	Improved accuracy
Image 6	Normal	56%	89%	Improved accuracy

In this project work, when the number is greater than 0.5, the risk is detected, which is abnormal and lesser than 0.5 means no risk detected and hence, it is a normal event. Around 75% of the training phase's results were accurate. The more of iterations will increase the model's accuracy. For testing purposes, the frames are taken from videos and kept in a single folder. Additionally, segments are made using each frame, and these segments are then compared with the inputs provided. The algorithm classifies the frames as either normal (running, walking) or suspicious (in public locations such road accident, assault, fighting, explosion, and shooting). A communication with the expected class will be forwarded to the appropriate authority in the event of abnormal behavior. The achieved accuracy is in the range of 85%. Its confusion matrix is seen in the figure 14 below. In general, the matrix is classified as:

a = [32, 11]

b = [8, 39]

Where, a is the tested negative (FP rate) and b is the tested positive (TP rate).

```
Confusion Matrix:
 [[32  8]
 [11 39]]
Accuracy 0.7888888888888889
```

Figure 14: Confusion Matrix

V. CONCLUSION AND FUTURE SCOPE

By keeping track of any suspicious activity near where security cameras are installed, the proposed project "Anomaly Detection in a Surveillance Video" will raise the bar for security. Monitoring of such bizarre incidents on several surveillance cameras using staff is prestigious. Hence, in this proposed system Anomaly or threatening events are detected and marked using Artificial Intelligence techniques. For detecting the anomalies in each segments(phases) of the video, the deep learning techniques has been utilized. An alert SMS will be sent about unusual events and upon receiving

the SMS siren will be turned on in the station. The abnormal scene discovery of reading from normal data is very important and unique applications. Also, the process of obtaining such anomalies depends entirely on the environment and unusual context.

The proposed algorithm architecture can be replaced with different ones such as with auto-encoders, with transformer model and other updated models in future. Further, one can monitor, detect and track down the various objects in videos or set of images by replacing with different suitable architecture, which is exceptional and convenient, can be used with the MIL algorithm. The accuracy and overall performance of the algorithm can further be improved by adding diverse datasets. Processing and detection phases can be improved further by using better algorithm. By doing such changes in the software, the result of the overall system can be upgraded for the future. The proposed system can further be enhanced to work in real-time by using micro-controller like Raspberry Pi, where live data collected from surveillance camera and in the particular data anomalies can be detected. Both the collection of data and anomaly detection can be done in one place and then alert to control station directly.

REFERENCES

[1] Amrutha C.V, C. Jyotsna, Amudha J, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video*", IEEE Xplore, ICIMIA 2020.*

[2] Waqas Sultani, Chen Chen, Mubarak Shah, "Real-world Anomaly Detection in Surveillance Videos", *IEEE 2018.*

[3] J. Wu, D. Zhang, Z. Zhou and Y. Li, "Anomaly object detection in urban management surveillance video based on deep learning", *2020 International Conference on Information Science and Education (ICISE-IE), 2020 IEEE, DOI: 10.1109/ICISE51755.2020.00017.*

[4] S. Saypadith and T. Onoye, "An Approach to Detect Anomaly in Video Using Deep Generative Network", *in IEEE Access, vol. 9, 2021.*

[5] Devashree R. Patrikar1 and Mayur Rajaram Parate, "Anomaly Detection using Edge Computing in Video Surveillance System: Review", *IIT Nagpur, research paper 2021.*

[6] G. S. R. Machiraju, K. A. Kumari and S. K. Sharif, "Object Detection and Tracking for Community Surveillance using Transfer Learning*", 6th International Conference on Inventive Computation Technologies (ICICT), 2021 IEEE, ISBN: 978-1-7281-8501-9.*

[7] Salima Omar, Asri Ngadi and Hamid H. Jebur, "Machine Learning Techniques for Anomaly Detection: An Overview*", ResearchGate 2013.*

[8] Website reference - The 5 Steps of Machine Learning (linkedin.com)

[9] Prerana Singh, "Data Leakage in Machine Learning: How it can be detected and minimize the risk", *Published in*

[13] Hung Vu, Tu Dinh Nguyen, Anthony Travers, Svetha Venkatesh and Dinh Phung, "EnergyBased Localized Anomaly Detection in Video Surveillance", Springer International Publishing AG, 2017.

[14] X. Ma et al., "A Comprehensive Survey on Graph Anomaly Detection with Deep Learning", in IEEE Transactions on Knowledge and Data Engineering, 2021.

[15] Tangqing Li, Zheng Wang, Siying Liu, and Wen-Yan Lin, "Deep Unsupervised Anomaly Detection", IEEE, WACV 2021.

[16] Wanting Zhang, Le Gao, Shaoyong Li and Wenqi Li, "Anomaly Detection with Partially Observed Anomaly Types", IEEE CCNS 2021.

[17] Prakhar Singh and Vinod Pankajakshan, "A Deep Learning Based Technique for Anomaly Detection in Surveillance Videos", IEEE NCC 2018.

[18] Imtiaz Ullah and Qusay H. Mahmoud, "Design and Development of a Deep Learning-Based Model for Anomaly Detection in IoT Networks", IEEE access 2021.

[19] Varun Chandola, Arindam Banerjee and Vipin Kumar, "Anomaly detection: A survey", ACM Journal, 2009.

[20] Weiming Hu, Tieniu Tan, Liang Wang and Steve Maybank, "A Survey on Visual Surveillance of Object Motion and Behaviors", IEEE 2004.

[21] P. Wu, J. Liu, and F. Shen, "A deep one-class neural network for anomalous event detection in complex scenes" , IEEE 2019.

[22] UCSD Anomaly Detection Dataset website papers.

[23] Transfer Learning basics and Understanding -website: https://machinelearningmastery.com/transfer-learnin-for-deep-learning/.

[24] Guansong Pang, Chunhua Shen, Longbing Cao and Anton van Hengel, " Deep Learning for Anomaly Detection: A Review ", ACM computing survey, 2021.

[25] Salima Omar, Asri Ngadi and Hamid H. Jebur, "Machine Learning Techniques for Anomaly Detection: An Overview", ResearchGate 2013.

[26] Mauro Di Pietro, "Deep Learning with Python: Neural Networks (complete tutorial)" - website: Towards Data Science, 2021.

[27] Raghavendra Chalapathy, Sanjay Chawla, "Deep Learning for Anomaly Detection: A Survey", ResearchGate 2019.