Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 2s (2024)

Real-Time Anomaly Detection in Video Surveillance: A Mathematical Modeling and Nonlinear Analysis Perspective with MobileNet and Bi-LSTM

Deepak Mane[1], Sheetal Phatangare[2], Siddhant Nawale[3], Siddhesh Wani[4], Varun Gujarathi[5] and Vaishnav Loya[6]

[1,2,3,4,5,6]Vishwakarma Institute of Technology Pune-411037, Maharashtra, India.

[1]dtmane@gmail.com

[2]sheetal.phatangare@vit.edu

[3]siddhant.nawale20@vit.edu

[4]siddhesh.wani20@vit.edu

[5]varun.gujrathi20@vit.edu

[6]vaishnav.loya20@vit.edu

Article History:

Received: *18-03-2024*

Revised: *10-05-2024*

Accepted: *26-05-2024*

Abstract:

For the main purpose of security and safety, video surveillance systems are everywhere and their need for presence is at the top. The prominent application of a video monitoring system is to track and monitor the live footage and recognize any unusual activity or behavior. For the prevention of any unwanted event or dangerous activity, it is mandatory to get early detection from video surveillance. The previous architectures of traditional systems were primarily based on manual monitoring anomaly detection which had increased the false negative rate of anomaly detection. In this paper, we proposed a customized framework that leverages MobileNet and Bi-LSTM for the accurate detection of anomalies in real-time with unsupervised learning techniques. Due to the spatial-temporal feature pattern and multi-modal functionality of MobileNet, it helps in increasing accuracy compared to previous systems. MobileNet uses depth-wise and point-wise convolution than the normal convolution which improves the performance metrics. This paper has illustrated a brief study of benchmark datasets such as UCF dataset including their performance in anomaly detection compared with traditional systems and existing deep learning approaches. The proposed model has given an accuracy of 95.33% compared to the latest year 2023 model which gave an accuracy of 90.80% on the same dataset.

Keywords: Accuracy; Anomaly detection; Bi-LSTM; Video surveillance; MobileNet; datasets; UCF dataset; Unsupervised learning techniques.

1. Introduction

Nowadays, crime and accident rates have drastically increased, and the domain of crime is variable, such as arson, assaults, arrests, road accidents, explosions, burglary, fighting, abuse, robbery, shooting, shoplifting, stealing, vandalism, etc. For detecting these anomalies, video monitoring systems are very important everywhere for early detection of anomalies and prevention of any hazardous event or unwanted activity. Traditional and previous systems were based on human interaction and manual control intervention. Any detection of unusual behavior or unwanted activity was done by manually monitoring and tracking all the live or past footage of video surveillance systems [3]. Processing and monitoring large data of video surveillance systems is a challenging and time-consuming process manually.

Due to this traditional system, the chances of missing any unusual activity from the sight of the supervisor of the surveillance system were high. Of this, any accidental anomaly could happen at any time. This unusual activity or abnormal behavior is known as an anomaly.

The author of this paper introduced a robust and better method than traditional systems, which successfully prevented and reduced the false negative rate of anomaly detection errors from video surveillance in real time. The advancements of these automated anomaly detection systems help supervisors of video surveillance systems to identify any anomaly without time-consuming processes with better accuracy. It alerts for taking immediate actions to effectively prevent any hazardous activity or behavior. The proposed system is based on the MobileNet model with a Bi-LSTM model for better accuracy and prediction of anomaly detection in real time. This proposed model can detect anomalies in real-time without requiring human intervention throughout the detection process. Applications of the proposed system include anomalies of various types, including arson, assaults, arrests, road accidents, explosions, burglary, fighting, abuse, robbery, shooting, shoplifting, stealing, vandalism, etc. can be easily identified by this automated proposed system, in real-time. As MobileNet has the ability to recognize temporal and spatial patterns, it enhances the compatibility of models with complex structures and pattern relationships, allowing them to be a suitable system for dynamic scenes using video surveillance systems. The system is solely based on unsupervised learning techniques as supervised learning techniques require labels, and they buffer when they are processed to handle large n-dimensional video data which results in an incapable to simulate the representation of the model.

The aim of this paper is

- To identify unusual patterns from video surveillance and predict the class of anomaly to which it belongs.
- To be able to detect anomalies in real-time from video monitoring systems.

The outline of this paper is as follows: Overall background along with literature survey of anomaly detection systems in section II, Proposed architecture in section III, Dataset analysis in section IV, Experimented Results in section V, Conclusion elaborated in section VI, and references are given in section VII.

2. Related Work

The main task of the video monitoring system is to detect outliers. Traditional methods of video monitoring systems require human input or human intervention. This system requires consistent human input. Video surveillance systems gather data 24/7. So, it becomes difficult for any human to stay consistent for any visual attention over longer periods. Over time by time, visual attention gets drastically reduced of any supervisor who is tracking and monitoring video surveillance system under his recognition. Many deep learning advanced techniques have been introduced past 10 years which do not rely on any human intervention and which detect anomalies in a consistent manner with minimal false negative rates. Research is still carried out on this topic as it is one of the complex and crucial topics in the domain of deep learning.

In this section of related work, the authors of this paper researched over 15 research papers and their articles. The best two papers of every year ranging from 2017 to 2023 are aligned sequentially year by year. So, all the techniques can be analyzed sequentially and their evolution year-wise.

In the first paper of the year 2017, authors used a two-phase background estimation module which mainly relies on a stable background. In every video, the first background will be depicted from the main frame and the movable foreground will only be captured as a main frame. As the background gets stable, it has no use for further processing. They used a deep learning classifier for augmented reality to depict the patches in line. The main challenges they faced were problems in resolving proper background estimation along with activity recognition [1]. In the paper on video anomaly detection of the year 2017 in which they proposed a system for human recognition through activity, faced the temporal correspondence throughout the frames. As temporal correspondence varies the data time by time. It becomes difficult to augment the data frame by frame. They used Kalman filtering and vector-based feature extraction for their proposed model which helped them track initial objects in a path named as eigen process[2]. In 2018, one paper proposed a system using CNN and ConvLSTM for enhanced temporal pattern classification with an accuracy of 74.8% and 80.2% for the UCSD Pedestrian 1 and UCSD Pedestrian 2 datasets. They faced challenges in their system regarding the intricate nature of anomalies in their contextual patterns along with broader adaptability [3]. Another paper of 2018 emprises the main challenges of online alerts and speed detection which gives them errors while training their model as their training model was taking much higher time compared to the expected time. They elaborated on the issue of localization and its absence in the existing systems. Localization tells the feature location of the anomaly present in the given frame. They diverted their architecture mainly into two phases as a train network and a detection classifier where there were five detection classes along with their scores. Each detection class will record its own detection score and the maximum detection score will be rewarded as the final detection score and then anomaly will be detected [4]. In the next year 2019, the first paper proposed a system including CNN as their main framework for detecting anomalies and SI and MHOF they had used as a spark ignition model along with tackling a few challenges associated with RGB levels and their edges and tracking of objects from angle derivations from multiple camera feeds. They impressed principal component analysis for obtaining a higher-graded feature [5]. In another paper of 2019, the proposed system was trained and tested on three different datasets as only one dataset was not efficient for checking the efficiency of the model. They used the CUHK dataset, UCSD Pedestrian Set 1, and UCSD Pedestrian Set 2 as their three major datasets. The main method was based on attention-driven loss which they segregated into two losses. They were facing problems in their model for handling and defining all negative samples [6]. In the year 2020, the authors of this paper used performance metrics at the very stage level for tuning and analyzing the efficiency of the system along they introduced eyeball evaluation as their new metrics and the main component of the model. It is the evaluation metrics of a detector that analyzes all the graphs and contains all the information. They faced the problem of lack of labeled data in a particular class as class imbalance and time factor for anomaly detection were comparatively higher than others [7].

Another 2020 paper described their model on the VGG-16 implemented via Imagenet dataset and they used UT interaction for their proposed algorithms named CNNs and LSTM. Their proposed model was only certain restricted to academic areas only [8]. The 2021 year papers mainly comprised new and efficient techniques as the first paper proposed their model in a set of 2 faster RCNN models and they did a classification process using deep reinforcement technique but they faced the problem in complexity while attaining normal activities and in the network layer. They used Faster CNNs for the reason of their high adaptability and their box regression coincidence [9]. Another paper focused on the deep learning method as a refined ResNet-50 model which was highly powerful and efficient. They did data augmentation for all the videos to be diverted from every orientation. They studied deeply the duty of law-enforcement agencies, their employees, and their challenges [10]. In the year 2022, this paper used HEVC which was formerly known as high-efficiency video coding for down sampling the videos for low compression to make them low bitrate from high bitrate. They had used ResNet50 and Bidirectional LSTM for their model [11]. Another paper focused on three main phases in their model as shot segmentation, spatio-temporal learning, and extraction of features along with LWCNN. However the primary ambiguity in their defined classes was high and there was an imbalance between positive and negative bitrates [12]. In 2023, this paper introduced an adaptive convolutional neural network as ACNN and they achieved an accuracy of 90.80% by taking three classes from the UCF dataset as fighting, road accidents, and explosion their three main classes. This was the latest paper by now which had implemented anomaly detection using deep learning techniques [13].

All Existing systems were based on three main parts such as supervised methods, semi-supervised methods, and unsupervised methods. The main problem associated with supervised methods was of labeling and its volumetric dimensional data. Data generated from video surveillance systems is dynamic and its environment changes from time to time, this term is called concept drift. This was the main liability of semi-supervised methods as they may not adapt well to process anomalies and detect them. Unsupervised methods were associated with the problem of higher computational complexities [4]. Most of the existing systems were running properly on small-scale datasets but eventually, they were giving a high percentage of false negative rate when implemented in real-time scenarios [12]. Traditional and existing systems rely on manual frame extraction which has increased the consumption of errors [11]. Other proposed systems in the past had implemented their models by using technologies like CNNs [3,5]. Some models had been implemented using ResNet50 [10,11] and a few models were based on RCNN architecture [9]. This proposed system was implemented on advanced CNN architecture which is MobileNet and Bi-directional Long short term memory (LSTM).

3. PROPOSED ARCHITECTURE

Existing systems faced the problem of high dimensional data size of the dataset containing more than 1000 videos set along with an overall size of more than 60 GB. The problem with the dataset was the videos were of high bitrate and high dimensional along with high space. The proposed system of the paper solved this problem. The authors of this paper used Adobe Media Encoder 2022 and VideoLAN Client 2023 for compressing the video set. Videos were taken individually from each class. The format was H.264, and the preset was match source – Adaptive low bitrate. The renderer used was mercury playback engine GPU Acceleration, which longs for CUDA

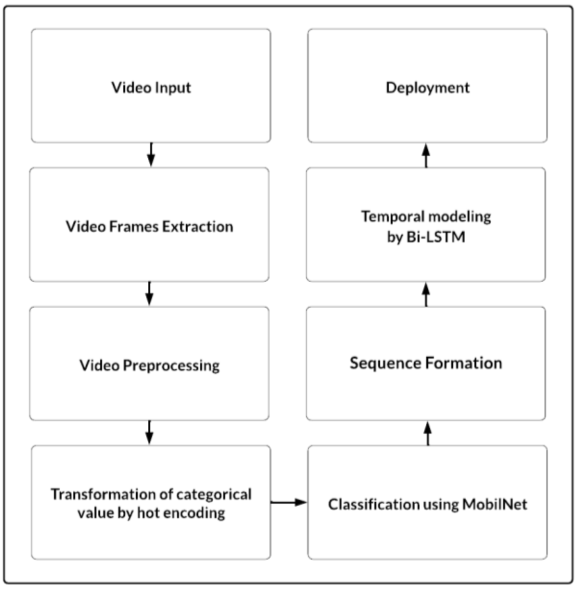

Figure 1: Proposed System Flow of MobileNet with Bi-LSTM

Figure 1 illustrates the proposed system flow of the proposed system. The predefined process starts with finding the proper utilized dataset for the model. Below are the sequence-wise proposed method of real-time anomaly detection from video surveillance using MobileNet and Bi-LSTM:

1.1. Video Input

Video input was the first primary step of the overall process method. The proposed model was evaluated on mainly three classes as Fighting, road accidents, and explosions from the UCF dataset [14]. So video was categorized according to the type of anomaly detected from its anomaly behavior by the system.

1.2. Video Frames Extraction

Extraction of video frames for each individual was carried out in the second step of the proposed method. The proposed system was totally based on the latest and advanced deep learning techniques such as MobileNet which is advanced CNNs architecture and Bi-LSTM(Bidirectional long short-term memory). Number of the keyframes per video depends on the overall size of the video capacity and its intensity along with the parameters included in the video. The size of the keyframes also depends on the video as more the size of the video, the more will be the size of the keyframes.

1.3. Video preprocessing

Video preprocessing was primarily focused on the removal of any unwanted objects and trimming of the unwanted clip which was taking unusual and unwanted space in the algorithm. The author of this paper used sequence length as a parameter that skips every other frame and holds the 16th frame subsequently in the algorithm.

1.4. Hot Encoding transformation

The primary data in the dataset was present in the form of categorical data. categorical data was used to gather all the same substantial feature patterns from all the keyframes extracted from the video. After the segregation of categorical data, it was converted into numerical data with the help of HOT encoding. All the categorical values were then converted into numerical data in the algorithm.

1.5. Classification using MobileNet

The main part of the system was classification using MobileNet. Here, the author of this paper used depth-wise convolution and point-wise convolution of MobileNet for the classification of feature patterns instead of normal convolution. Normal convolution works only standalone but the novelty of MobileNet architecture is their depth-wise convolution and point-wise convolution that seggregates all the input form depth-wise by sequentially going at each channel individually.

1.6. Sequence Formation

Sequence length is the total count per step which will directed to the bidirectional LSTM network. If we have any word embeddings then we can't directly direct those embeddings into bi-LSTM. For conversion of word to number embeddings, encoding the word embedding is a must and that's where sequence length is primarily used. Bi-directional LSTM requires the required number of sequence input of time steps. It will processed further into the parser and then it will diverge into three defined classes. The author of this paper used sequence length as a parameter that skips every other frame and holds the 16th frame subsequently in the algorithm.

1.7. Temporal modeling by Bi-LSTM

Temporal modeling in deep learning primarily focuses on the data that consistently varies over time period. Temporal modeling was attached with every video input on their video keyframes which were varying over the video duration in the form of seconds. Recurrent Neural Networks geanerally used for remembering the past consequences pattern. But if the pattern consequences are large and time-taking then bidirectional long short term memory is the recommended option and the author of this paper used the same.

1.8. Deployment of the detection system in real-time

The final step of the model was to deploy the anomaly detection system from video surveillance in real-time with the input format as video surveillance which can constantly give input for detection of any anomaly behavior from the live footage. The main deployment software used for the system was Google Colab.

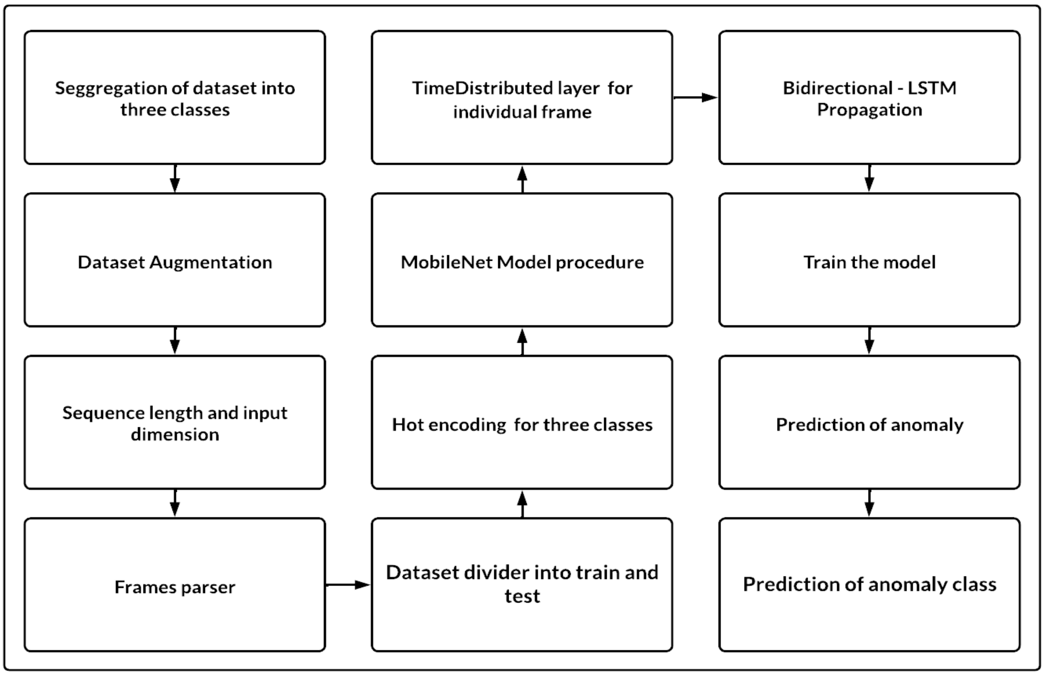

Figure 2: Module-wise flow of MobileNet and Bi-LSTM

The module-wise flow in figure 2 of the proposed model contains four major premises such as segregation and augmentation of dataset, sequence length and input dimension, frame parser, and prediction of anomaly. All the sections are described below sequentially:

2.1. Segregation and Augmentation of Dataset

UCF Dataset was taken for the evaluation and training of the proposed model. The three classes taken for the proposed model were fighting, road accidents, and explosions. The proposed model will also identify normal or abnormal behavior of the video input. For considering all the aspects and possibilities of anomaly, augmentation Was been performed on the whole dataset of all three classes.

The total count of processed videos of the whole dataset after augmentation is given below:

Table 1. Anomaly classes with count

Labels	Classes	Count
0	Explosion	500
1	Fightings	250
2	Road Accidents	500
3	Normal	1360
TOTAL		2610

The methods for augmentation that have been done on the dataset were horizontal flip, shearing of videos, brightness adjustment, and vertical flipping. In shearing, all the perception angles were involved in image distortion. Adjusting brightness will help for realistic and simulating some real-life situations.

2.2. Sequence length and input dimension

Sequence length is the total count per step which will directed to the bidirectional LSTM network. If we have any word embeddings then we can't directly direct those embeddings into bi-LSTM. For conversion of word to number embeddings, encoding the word embedding is a must and that's where sequence length is primarily used. Bi-directional LSTM requires the required number of sequence input of time steps. It will processed further into the parser and then it will diverge into three defined classes.

2.3. Frames Parser

Frames parser contains some predefined processes related to trimming, adjusting, and modeling of frames along with the extraction of frames as a main component. Pre-defined processes include analyzing current and latest frames and appending all the latest frames in the list along with the main paths of video input destinations. All the frames and video input data was balanced by adjusting their frame size as all the data of the dataset may not be balanced and of the same size and adjustments. Further, the dataset was split into training and testing datasets.

2.4. MobileNet Procedure

MobileNet is the main component of this proposed system of real-time anomaly detection as MobileNet is an advanced CNN architecture along with some advanced features that help the model to achieve the predicted accuracy with minimal false negative rate. The main novelty of MobileNet is its convolution framework as MobileNet works on depth-wise and point-wise convolutional networks. MobileNet is a lightweight architecture and follows streamlined architecture. Due to this, it also works sufficiently on mobiles and embedded devices. Hot encoding will primarily used for converting categorical data into numerical data. For example, three classes named fighting, road accidents, and explosion will converted into (0,1,2).

The coefficient equation of MobileNet as follows:

$$ar = \frac{\sum_{i=1}^{n} \left(\frac{a_i - a}{si_x}\right)\left(\frac{b_i - b}{si_y}\right)}{mn-1}$$

Equation 1: Coefficient equation of MobileNet

Equation 1 represents coefficient representation of MobileNet whereas ar stands for the output of the predictive data segments. In the above equation, a and b are variables that are the cardinal set for the variables a and b. ar is the derivation cross sum of the variables. si_x and si_y are the input sequence related with variables a and b which are associated with the final outcome ar.

2.5. Bi-LSTM

Bidirectional Long short-term memory is a sequential model that carries two inputs for its subsequent model as one input will processed for the forward network and one input will processed for the backward network. It is different from ConvLSTM which is convolutional long short term memory. Bi-directional LSTM supports two-way flow when compared to uni-directional LSTM. Bidirectional LSTM enhances its mechanism by adding one notch layer that supports and is able to reverse data flow from its centric system.

The formulation of Bi-LSTM as follows:

$$D_{ts}^{si} = tanh(Vh_{g} * h_{ts-1}^{si} + V_{ex}^{si} * e_{ts}^{si})$$

Equation 2: Formulational equation for Bi-LSTM

In the above Equation 2, Vh_g and V_{ex}^{si} represent volumetric weights associated with the model. si in the superscript of the equation means information related to the input sequence and ts in the subscript of the equation means the information related to the time sequence. The term D_{ts}^{si} here represents the cumulative output at a particular step. V_{ex}^{si} this is the other volumetric weight related to the output of particular time step.

Bi-directional LSTM was used to track and capture the temporal correspondence across the model. It varies over time. Timedistributedlayer was been used for each frame extraction process and the application of MobileNet model to each individual frame by frame.

The overall predictive MobileNet classifier of the proposed model consisting of total parameters, output shape listing, and the sequential layers. The layers included in this model are the time_distributed layer, dropout layer, bidirectional layer, and dense layer. This sequential model will process the same video data input from the time_distributed layer to the dense layer sequentially. The bidirectional layer is primarily based on LSTM.

2.6. Final prediction and predictive output

ADMM is an Alternating direction method of multipliers, batch gradient descent, stochastic gradient descent, evolutionary algorithms, and Adam optimizers These are some of the optimization techniques generally used for the predictive modeling of the models. Here, the most commonly and enhancive used method was the stochastic gradient descent method for optimizing the proposed model. The model was trained on the count of 100 epochs with a dropout of 0.5 and batch size 16. After completion of model training, it was been tested on the basis of all three classes along with normal and abnormal behavior detection classifiers.

3. Dataset

Most of the existing systems used the UCSD dataset for anomaly detection from video surveillance. UCSD dataset is divided into two categories as UCSD Pedestrian 1 and UCSD Pedestrian 2 [3,9]. Some of the models also used the CUHK dataset [6]. The proposed system was trained and tested on the UCF anomaly dataset which consists of a total of 128 hours of watch time. The reference link of the same dataset can be accessed easily at[16]. Videos of the dataset were untrimmed and sourced from real-time video surveillance. The dataset consisted of a total of 13 distinct anomaly classes such as arson, assaults, arrests, road accidents, explosions, burglary, fighting, abuse, robbery, shooting, shoplifting, stealing, vandalism, etc.

4. Results

Video inputs were processed correctly and the proposed model depicts abnormal behavior from normal behavior by showing the type of anomaly on screen in real-time with an accuracy of 95.33%. The model was trained for an overall 100 epochs which was a higher metric than most of the existing systems. The batch size was 16 along with a dropout of 0.5. All the performance and analytics results were been carried out and illustrated as follows:

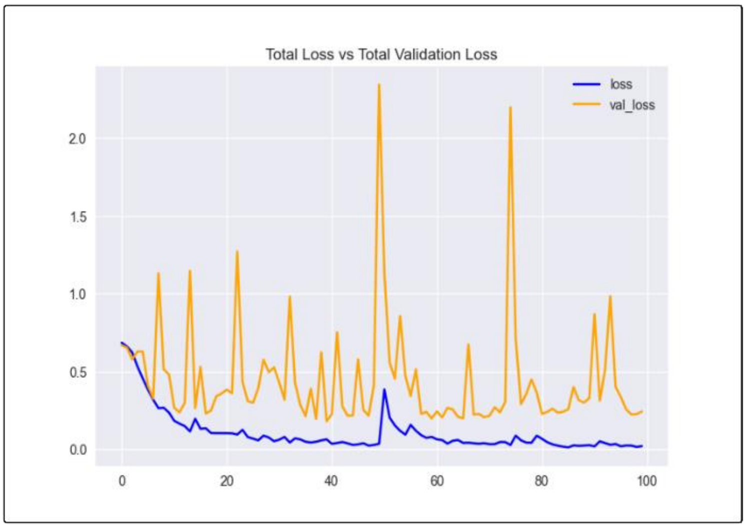

Figure 3: Loss graph for Normal vs. abnormal)

Figure 3 shows the accurate graph of total loss vs total validation loss compared between normal and abnormal behavior.

Figure 4: Accuracy graph for Normal vs. abnormal

Figure 4 shows the accurate graph of total accuracy vs. total validation accuracy compared between normal and abnormal behavior.

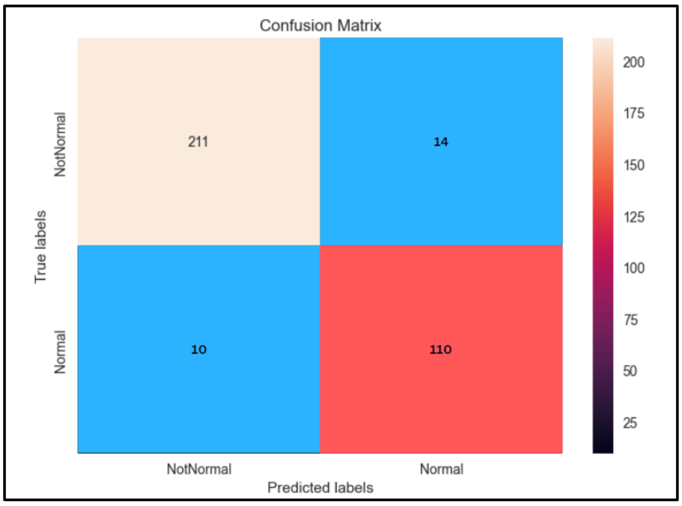

Figure 5: Confusion matrix for Normal vs. abnormal

Figure 5 shows the confusion matrix made for the normal and abnormal behavior of the video input which is the accurate prediction of each entity.

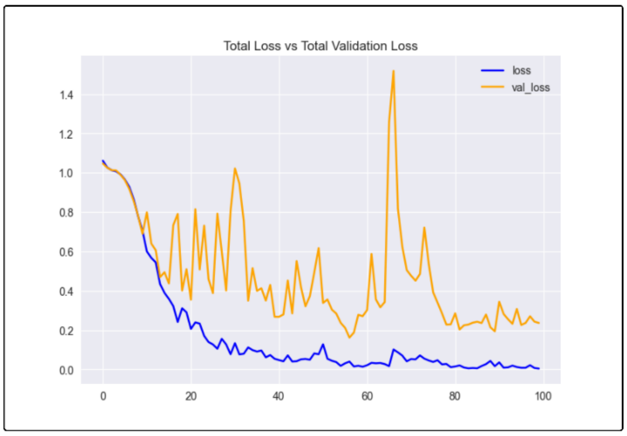

Figure 6: Loss graph for classes of Anomaly

Figure 6 shows the accurate graph of total loss vs. total validation loss for the defined three classes: fighting, road accidents, and explosion.

Figure 7: Accuracy graph for classes of Anomaly

Figure 7 shows the accurate graph of total accuracy vs. total validation accuracy for the defined three classes named as fighting, road accidents, and explosion.

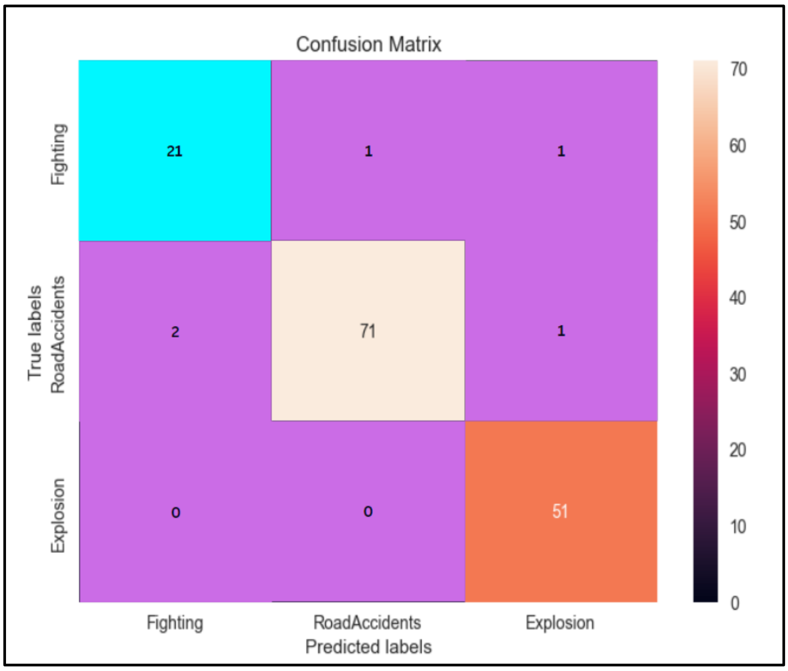

Figure 8: Confusion matrix for classes of Anomaly

Figure 8 shows the confusion matrix for the defined three classes named as fighting, road accidents, and explosion which is the accurate prediction of each entity.

Table 2: Performance metrics of the existing model

Class	Precision	Recall	f-score	Support
Anomaly	90.23	91.58	90.54	25,584
Non-anomaly	91.45	90.02	90.96	36,897

Table 3: Performance metrics of the proposed model

Class	Precision	Recall	f-score	Support
Anomaly	95.23	94.86	95.78	225
Non-anomaly	89.65	90.45	90.56	120

In Table 2 and Table 3, metrics evaluation of the existing model for the year 2023 and the proposed model of 2023 have been compared. Metrics compared here are precision, recall, f-core, and support. By comparing and analyzing both metrics, it has been clearly seen that the metrics performance of the proposed model was better than the existing model.

Table 4: Output Comparison with Existing Models

Reference	Algorithm Used	Accuracy
2017[1]	Kernel fuzzy C-Means	72.64%
2018[3]	CNN + ConvLSTM	74.82%
2018[4]	MCCNN	89.40%
2019[6]	Attention-driven loss	83.90%
2020[7]	CNN + LSTM	87.15%
2021[10]	ResNet-50	79.69%
2022[11]	(HEVC)-H265	90.16%
2022[12]	CNN + LSTM	93.00%
2023[13]	ACNN	90.80%
Proposed Model	**MobileNet+ Bi-LSTM**	**95.33%**

Table 4 shows the predictive comparison of the results of the existing model with the proposed model of this paper. Models were taken sequentially year-wise for comparison. It was clearly seen from the outcome that the accuracy of existing models ranged from 70% to 92% on average. The final outcome of the proposed model shows that the accuracy of the proposed system was high compared to the existing system by using MobileNet and bi-directional long short term memory.

4. CONCLUSION

The current proposed model of this paper showed accurate results and predictions compared to the existing system, and most of the existing systems were not fully engaged with real-time systems. The current model was engaged with real-time devices, and the false negative rate of the current model was minimal. The selected deep learning architectures, such as MobileNet and Bi-directional Long short-term memory, were perfectly suitable for predicting anomalies in real-time from video surveillance. The proposed model will surely add one layer of security patch to public safety.Further, the proposed model can be implemented on other datasets for analyzing the overall performance throughout the datasets. More classes can be added for anomaly detection in real time. So, the applications and the adaptability of the proposed system will be enhanced. The accuracy of the current proposed model can be improved with the new and latest deep learning architecture, and more distinct classes can be added.

REFERENCES

[1] Revathi, A. & Kumar, Dhananjay. (2017). An efficient system for anomaly detection using deep learning classifier. Signal, Image and Video Processing. 11. 10.1007/s11760-016-0935-0.

[2] S. Zaidi, B. Jagadeesh, K. V. Sudheesh and A. A. Audre, "Video Anomaly Detection and Classification for Human Activity Recognition," 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, India, 2017, pp. 544-548, doi: 10.1109/CTCEEC.2017.8455012.

[3] P. Singh and V. Pankajakshan, "A Deep Learning Based Technique for Anomaly Detection in Surveillance Videos," 2018 Twenty Fourth National Conference on Communications (NCC), Hyderabad, India, 2018, pp. 1-6, doi: 10.1109/NCC.2018.8599969.

[4] A. Khaleghi and M. S. Moin, "Improved anomaly detection in surveillance videos based on a deep learning method," 2018 8th Conference of AI & Robotics and 10th RoboCup Iranopen International Symposium (IRANOPEN), Qazvin, Iran, 2018, pp. 73-81, doi: 10.1109/RIOS.2018.8406634.

[5] A. Mohan, M. Choksi and M. A. Zaveri, "Anomaly and Activity Recognition Using Machine Learning Approach for Video Based Surveillance," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-6, doi: 10.1109/ICCCNT45670.2019.8944396.

[6] J. T. Zhou, L. Zhang, Z. Fang, J. Du, X. Peng and Y. Xiao, "Attention-Driven Loss for Anomaly Detection in Video Surveillance," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 12, pp. 4639-4647, Dec. 2020, doi: 10.1109/TCSVT.2019.2962229.

[7] R. J. Franklin, Mohana and V. Dabbagol, "Anomaly Detection in Videos for Video Surveillance Applications using Neural Networks," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2020, pp. 632-637, doi: 10.1109/ICISC47916.2020.9171212.

[8] C. V. Amrutha, C. Jyotsna and J. Amudha, "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video," 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 2020, pp. 335-339, doi: 10.1109/ICIMIA48430.2020.9074920.

[9] Romany F. Mansour, José Escorcia-Gutierrez, Margarita Gamarra, Jair A. Villanueva, Nallig Leal,

[10] “Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model”,Image and Vision Computing,Volume 112,2021,104229,ISSN 0262-8856,https://doi.org/10.1016/j.imavis.2021.104229.

[11] Rahman, Md & Afrin, Mst & Atikuzzaman, Md & Rahaman, Muhammad. (2021). Real-Time Anomaly Detection and Classification from Surveillance Cameras Using Deep Neural Network. 1-6. 10.1109/STI53101.2021.9732584.

[12] Abbas, Zainab & Al-Ani, Ayad. (2022). Anomaly detection in surveillance videos based on H265 and deep learning. International Journal of Advanced Technology and Engineering Exploration.9.910-922. 10.19101/IJATEE.2021.875907.

[13] Ul Amin, S.; Ullah, M.; Sajjad, M.; Cheikh, F.A.; Hijji, M.; Hijji, A.; Muhammad, K. EADN: An Efficient Deep Learning Model for Anomaly Detection in Videos. *Mathematics* **2022**, *10*, 1555. https://doi.org/10.3390/math10091555

[14] Mane, D., Kumbharkar, P., Pawar, P., Katkar, K., Shah, S., Jamwal, K. (2023). Anomaly Detection from Video Surveillances Using Adaptive Convolutional Neural Network. In: Manchuri, A.R., Marla, D., Rao, V.V. (eds) Intelligent Manufacturing and Energy Sustainability. Smart Innovation, Systems and Technologies, vol 334. Springer, Singapore. https://doi.org/10.1007/978-981-19-8497-6_21

[15] Mane, Deepak & Sangve, Sunil & Upadhye, Gopal & Kandhare, Sahil & Mohole, Saurabh & Sonar, Sanket & Tupare, Satej. (2022). Detection of Anomaly using Machine Learning: A Comprehensive Survey. International Journal of Emerging Technology and Advanced Engineering. 12. 134-152. 10.46338/ijetae1122_15.

[16] [Minhajuddinmeraj]. ([2022, October]). [Anomaly-Detection-Dataset-UCF], [Version: 01]. from [https://www.kaggle.com/datasets/minhajuddinmeraj/anomalydetectiondatasetucf].

[17] Dhabliya, Dharmesh , Kunche, Shanthi , Dingankar, Shreyas , Dari, Sukhvinder Singh , Dhabliya, Ritika & Khetani, Vinit (2024) Blockchain technology as a paradigm for enhancing cyber security in distributed systems, Journal of Discrete Mathematical Sciences and Cryptography, 27:2-B, 729–740, DOI: 10.47974/JDMSC-1923

[18] Goyal, Dinesh , Kumar, Anil , Gandhi, Yatin & Khetani, Vinit (2024) Securing wireless sensor networks with novel hybrid lightweight cryptographic protocols, Journal of Discrete Mathematical Sciences and Cryptography, 27:2-B, 703–714, DOI: 10.47974/JDMSC-1921

[19] Toumlilin, M., Abidin, M.Z., "Well-posedness of the 3D Stochastic Generalized Rotating Magnetohydrodynamics Equations" (2022) Advances in the Theory of Nonlinear Analysis and its Applications, 6 (4), pp. 513-527.