International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

VADNet: A Novel Deep Learning Architecture for Automatic Detection and Classification of Abnormalities from Public Surveillance Videos

[1]Ganta Raju, [2]Gs Naveen Kumar,

Submitted: 28/01/2024 **Revised**: 06/03/2024 **Accepted**: 14/03/2024

Abstract: Usage of public surveillance cameras in cities and other places of public importance has resulted in reduction of crimes besides helping in establishing evidences of certain incidents. Thus, video analytics has become indispensable research area. With technological innovations like cloud computing, Internet of Things (IoT) and Artificial Intelligence (AI), storage and analysis of videos in real time is made possible. At the same time, emerging techniques in deep learning have paved way for processing of image content leading efficient video analytics. Abnormality detection from surveillance videos has assumed significance. The existing research in this area based on Convolutional Neural Network (CNN) has limitations because of the fact that it cannot bestow optimal performance unless, it is customized to solve the problem in hand. Towards this end, in this paper, we proposed a novel deep learning architecture known as Video Abnormality Detection Net (VADNet) for detection of abnormalities from surveillance videos. VADNet is a CNN variant designed for leveraging detection performance. We proposed an algorithm named Learning based Video Abnormality Detection (LbVAD) which exploits VADNet for efficient detection of video abnormalities. UCF-Crime is the benchmark dataset used for our empirical study. Our experimental results revealed that VADNet outperforms existing CNN variants like MobileNetV1, ResNet50 and VGG19 models with highest accuracy 95.64%.

Keywords: *Video Abnormality Detection, Public Surveillance, Deep Learning, Artificial Intelligence, Novel Deep Learning Architecture*

1. Introduction

Surveillance videos, of late, used in cities could improve public safety by reducing crime rate besides helping in faster convergence in different investigations. Public video surveillance has its positive impact on the society. The usage of deep learning models for video analytics has paved way for automatic detection of abnormalities from such videos [1]. Video content analysis can be done in two ways. First, it is done when there is some incident occurs, which is not real time. This approach is time consuming and prone to giving chance to offenders to escape. Second approach is automated approach using AI which functions in real time and provides notifications to authorities concerned [2]. In the presence of thousands of surveillance videos producing video content, it is desirable to have such automated approach for detection and classification of abnormalities in real time. There are many contributions in this regard. Many models, as discussed in [4], [5] and [6], are based on CNN as it could improve image processing performance. Rezaee et al. [4] explored effectively identifying anomalous actions in congested environments that improves public safety. Nayak et al. [5] found that for safety purposes, video monitoring is frequently used in public spaces. Khan et al. [6] addresses issues like accident detection, particularly in traffic by Video surveillance.

[1]*Research Scholar, Department of CSE, MALLAREDDY UNIVERSITY, HYDERABAD, TELANGANA.*
[2]*Associate Professor Department of Data science & Information Technology MALLAREDDY UNIVERSITY, HYDERABAD, TELANGANA*
Email: [2]gsrinivasanaveen@gmail.com
Corresponding Author: rishi.shinny@gmail.com

This work achieves 82% accuracy in anomaly detection by using CNNs. A hybrid approach is found in [35] where CNN and SVM models are used for abnormality detection. There are approaches that exploited generative adversarial network (GAN) architecture also as discussed in [36]. Traffic surveillance for abnormality detection is explored in [6], [29] and [30]. In essence, deep learning models are widely used for video analytics. From the review of literature, it was observed that CNN models used for video anomaly detection need enhancements for leveraging performance. Our contributions in this paper as follows.

1. We proposed a novel deep learning architecture known as Video Abnormality Detection Net (VADNet) for detection of abnormalities from surveillance videos.
2. We proposed an algorithm named Learning based Video Abnormality Detection (LbVAD) which exploits VADNet for efficient detection of video abnormalities.
3. A prototype is built to evaluate VADNet using a benchmark dataset known as UCF-Crime dataset.
4. Our experimental results of VADNet are compared with existing CNN variants like MobileNetV1, ResNet50 and VGG19.

The remainder of the paper is structured as follows. Section 2 reviews literature on different models existing for video abnormality detection. Section 3 presents our methodology and deep learning architecture. Section 4 presents results of our empirical study. Section 5 provides discussion and limitations of the study. Section 6 concludes our model and provides scope for future research.

2. Related Work

This section reviews literature on existing deep model used for video abnormality detection. Khaleghi and Moin [1] observed that increased use of surveillance cameras in recent decades has aided in anomaly identification and event prediction. Deep learning greatly improves accuracy, particularly with convolutional neural networks. Nawaratne et al. [2] investigated on the need for intelligent real-time video monitoring that is driven by autonomous industries and urbanization. Updating anomaly detection in changing behaviours, the proposed ISTL tackles issues. Kavikuil and Amudha [3] found that because of shifting conditions and people moving about, it might be difficult to identify abnormalities in video surveillance. Accuracy in feature learning is enhanced by deep learning's efficiency. Rezaee et al. [4] explored effectively identifying anomalous actions in congested environments that improves public safety. Detection is improved by real-time monitoring employing machine learning and the WoT platform. Nayak et al. [5] found that for safety purposes, video monitoring is frequently used in public spaces. Despite obstacles, deep learning improves anomaly detection.

Khan et al. [6] addresses issues like accident detection, particularly in traffic by Video surveillance. This work achieves 82% accuracy in anomaly detection by using CNNs. Munyua et al. [7] observed that in computer vision, deep learning is essential, particularly for anomaly detection in security footage. This evaluation benchmarks deep learning solutions since 2016. Aberkane and Elarbi [8] studied surveillance films used in smart cities; a Deep Q Learning Network (DQN) uses deep learning and reinforcement methods to find abnormalities in the UCF dataset. Jebur et al. [9] observed that computer vision now places a strong emphasis on detecting human behaviour thanks to advances in technology. Deep learning is used to solve anomaly detection problems. Cosar et al. [10] suggested to use pixel analysis and trajectory analysis together for unified aberrant behaviour identification. By recognizing a variety of anomalous group behaviours, the technique decreases false alarms.

Giannakeris et al. [11] explored trends in autonomous driving and a framework for traffic analysis is created. Vehicle speed estimate and anomalous event detection are improved by cooperative detection and tracking using deep CNN features. Balasundaram et al. [12] opined that in a variety of settings, video surveillance is essential for anomaly detection. Real-time anomaly detection is improved by the use of pixel-wise frame displacement in a unique block-based technique. Ahmed et al. [13] observed that the use of CCTV in surveillance is growing. This research suggests a graph-based method for expert surveillance systems that combines information to identify anomalies in trajectory. Lin et al. [14] presented a technique that improves detection accuracy by identifying unusual activity in surveillance footage and producing precise summary movies. Meng et al. [15] presented a unique spatio-temporal deep network-based abnormal event identification technique that successfully solves the problem of limited aberrant samples.

Asad et al. [16] investigated and found that for the sake of public safety, automatic identification of unusual activity in surveillance footage is essential. An innovative two-stage design greatly enhances performance. Ye et al. [17] proposed a hybrid modulation technique for abnormal event detection that uses feature expectation subgraph calibration to increase classification accuracy. Ko and Sim [18] ensured speed and accuracy in the instantaneous identification of anomalous human behaviour in video surveillance, a unified framework utilizing deep convolutional structures is presented. Prawiro et al. [19] used an auto encoder to learn typical patterns and evaluating irregularities through reconstruction mistakes, one may identify abnormal occurrences in surveillance films. We suggest a two-stream decoder model for this job. Li et al. [20] presented a novel method for detecting abnormal events in crowded scenes by using a dictionary learning algorithm with compact and low rank.

Paul et al. [21] examined ways for reliably identifying people in surveillance footage, including object identification and categorization utilizing a range of approaches and databases. Mabrouk and Zagrouba [22] examined feature extraction, classification techniques, datasets, assessment metrics, and behaviour representation and modelling in video surveillance systems. Yan et al. [23] presented a two-stream structure with recurrent variation auto encoders for appearance and motion characteristics as part of a semi-supervised learning approach for abnormal event identification in video surveillance. Results from experiments show that it works well on several benchmark datasets. Wang et al. [24] showed better results on benchmark datasets when applying a two-stage method with an auto encoder and a one-class SVM for unsupervised anomalous event identification. Chong et al. [25] used a spatiotemporal architecture, the suggested approach for anomaly identification in movies achieves rapid, high-accuracy results that are on par with state-of-the-art techniques.

Li et al. [26] presented a Two-Stream DSTAE model that uses temporal and spatial streams to extract attributes in order to accurately identify video anomalies. Leyva et al. [27] identified the shortcomings in video anomaly detection testing and suggests using the Live Videos (LV) dataset, which consists of authentic surveillance footage, for assessment. Zhou et al. [28] suggested combining an LSTM encoder-decoder with a hybrid auto encoder to detect anomalies in videos. This would increase the decoder's extrapolation capabilities and spatiotemporal context, leading to better results on benchmark datasets. Cui et al. [29] found that in order to identify areas based on behaviour for early alerts in Intelligent Traffic Systems (ITS), a traffic video surveillance system that employs local attributes to detect anomalous occurrences is proposed. Dong et al. [30] explored a unique technique that uses directional motion behaviour descriptors without pre-set zones to automatically detect traffic abnormalities in busy scenarios.

Leyva et al. [31] presented a binary-based system that emphasizes faster processing speeds and more accuracy than previous online techniques for detecting aberrant events in internet videos. Varadarajan et al. [32] suggested by utilizing pLSA to analyse complicated actions in surveillance data in an unsupervised manner, allowing for the segmentation of scenes based on semantics and the identification of anomalies. Leyva et al. [33] improved performance on many datasets by introducing an online video anomaly detection system that uses binary features for motion encoding and low-complexity probabilistic models. Sharma et al. [34] focused on structural distortion-based abnormality scoring in its spatiotemporal model for anomaly identification in security footage. The paper also presents a defensive mechanism and investigates adversarial assaults on abnormality detection. Rajendran and Sungeetha [35] tackled the problem of storage optimization and dimensionality reduction in surveillance photos for effective video anomaly detection. The suggested method compares several current methodologies and uses CNN and SVM for classification.

Ravanbakhsh et al. [36] suggested by employing Generative Adversarial Nets (GANs) to detect abnormalities in crowded environments, demonstrating improved performance across many datasets for anomaly identification. Zhang et al. [37] presented a method for using blob sequence optimization to find anomalous activity in surveillance footage and create appropriate summary movies. Salim et al. [38] focused on accuracy and flexibility over sensor-based or human-based solutions by introducing a video-based framework for crowd recognition, tracking, and counting. Popoola et al. [39] examined approaches to modelling human behaviour, concentrating on the identification of anomalous behaviour in applications related to video surveillance and tackling practical issues. Karim and Shati [40] proposed a method by using the information from interest point pairs in two datasets to apply k-Means clustering for aberrant detection, yielding varying results and suggesting opportunities for improvement. From the review of literature, it was observed that CNN models used for video anomaly detection need enhancements for leveraging performance.

3. Propoased System

This section presents the proposed system and our deep learning architecture known as VADNet for video abnormality detection. It throws light on different aspects of the research carried out besides layers involved in the proposed deep learning architecture.

3.1 Problem Definition

Provided a streaming video collected from surveillance camera, development of a novel deep learning architecture for automatic detection of abnormalities is the challenging problem considered.

3.2 Our Framework

We proposed framework which illustrates the methodology for automatic detection of abnormalities from surveillance videos. Our methodology is based on deep learning. We proposed a novel deep learning architecture, a CNN variant, for leveraging detection performance. Thus, our proposal included a unique CNN-based deep learning architecture called VADNet, along with a methodology for effectively identifying abnormalities in surveillance footage. Prior to delving into the technical aspects of VADNet, we explore its overarching methodology, which offers a detailed explanation of our approach's mode of operation. Our framework, shown in Figure 1, takes UCF-Crime dataset [41] as input and follows a supervised learning method for detection and classification of abnormalities. It splits the dataset into 75% and 25% data towards training and testing respectively. The data is then subjected to analysis to know whether there is need for data augmentation. To improve quality of training data augmentation techniques such as flipping, shifting, zoom and rotation range are used. A deep learning model is built for training and performing abnormality detection. The proposed deep learning model VADNet is illustrated in Figure 2.

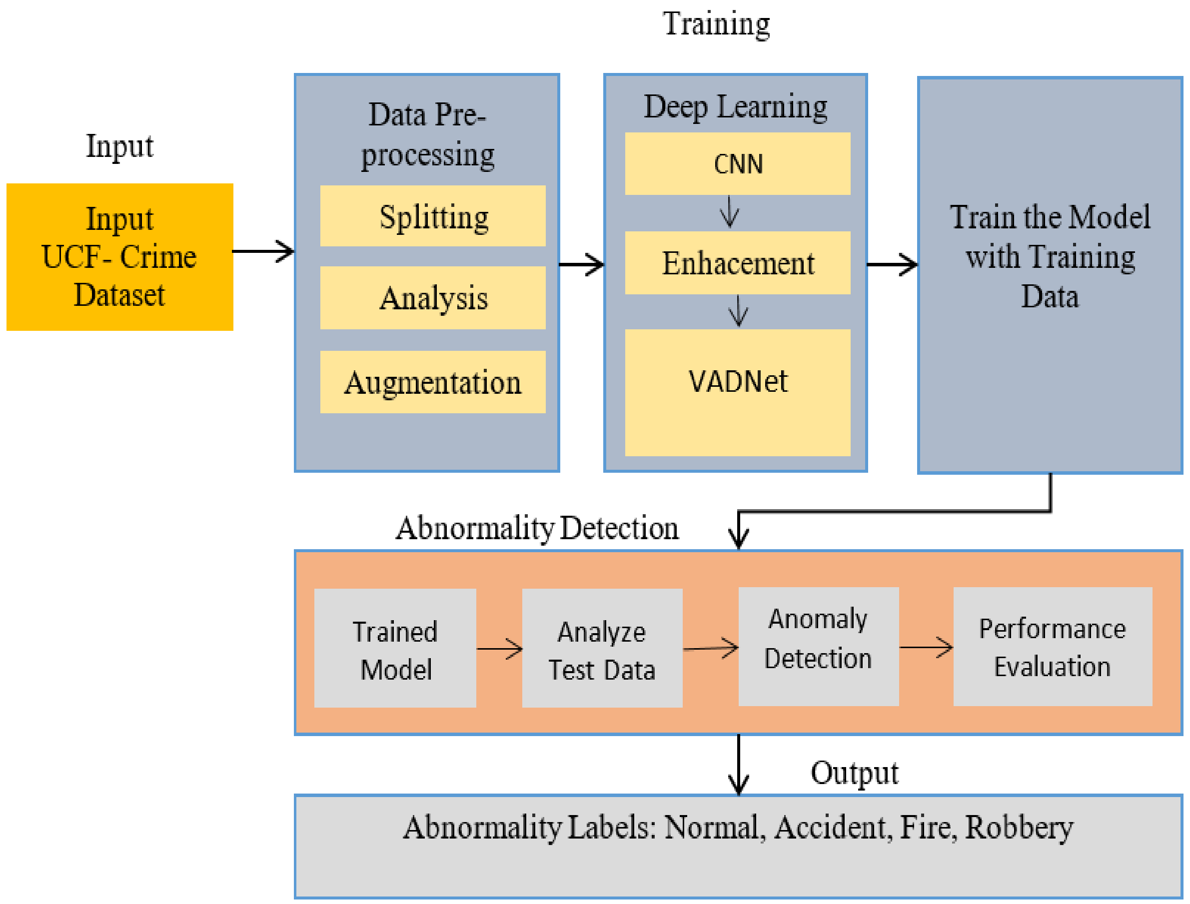


Fig 1: Our framework for video abnormality detection

Once the VADNet is created, it is compiled and the model is used for training using the 75% training data. Once the training is completed, the model is persisted for future reuse. Then the model is used for abnormality detection from given test videos. There is provision for detection and also classification of abnormalities. The training data has four classes such as normal, robbery, fire and accident. The model is designed to work with any number of abnormality classes, provided the training data, with appropriate softmax classifier for multi-class classification.

3.3 Novel Deep Learning Architecture

We proposed a novel deep learning architecture, a CNN variant, known as VADNet. This model has multiple layers and the fully connected layers followed by softmax activation function for abnormality classification. Figure 2 presents architecture of VADNet with different layers involved.

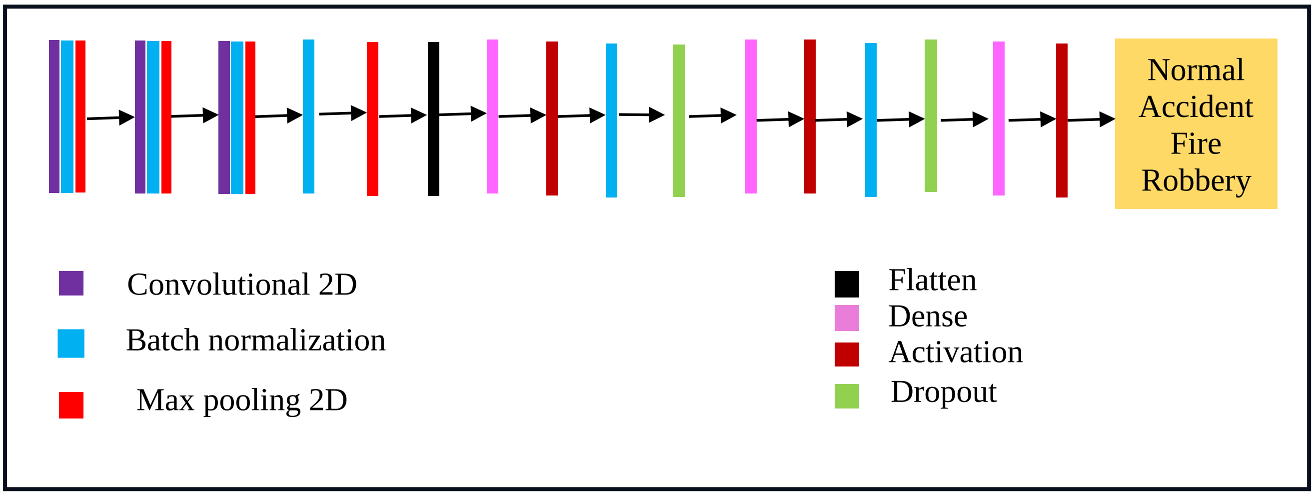


Fig 2: Proposed deep learning architecture named VADNet

As presented in Figure 2, our empirical research was aimed at building the VADNet architecture, which maximizes performance in anomaly identification from surveillance footage. There are several levels in our model. These observations on the strata are provided. A separable convolutional 2D layer with 16 filters is applied initially. Batch normalization and max pooling 2D with pool size (2,2) come next. This three-layer combination is then replicated, except this time the separable convolutional 2D layer has 32 filters. The next step is to add batch normalization and a separable convolutional 2D with 64 filters twice. This is followed by a max pooling 2D layer with a pool size of (2,2). Two sets of fully connected layers are added thereafter. Flatten layer, dense layer, batch normalization layer, ReLU activation layer, and dropout layer make up the first set. The ReLU activation layer, batch normalization layer, dropout layer, and dense layer make up the second set. A softmax layer is placed after these two fully connected layers to do the final multi-class classification. The activation layer comes after the dense layer to form the softmax layer. Table 1 shows the summary of layers of VADNet model.

Layer	**Convolution**	**Batch Normalization**	**Max Pooling**		
First layer	16 filters, (7,7) kernel size, “same” padding, ReLU activation	chanDim axis	(2,2) pool size, (2,2) strides		
Second Layer	32 filters, (3,3) kernel size, “same” padding, ReLU activation	chanDim axis	(2,2) pool size, (2,2) strides		
Third Layer	(64 filters, (3,3) kernel size, “same” padding, ReLU activation) x2	chanDim axis	(2,2) pool size, (2,2) strides		
Fourth Layer (First FC layer)	**Flatten**	**Dense**	**Activation**	**Batch Normalization**	**Dropout**
	Yes	128	ReLU	Yes	0.5
Fifth Layer (First FC layer)	**Flatten**	**Dense**	**Activation**	**Batch Normalization**	**Dropout**
	-	128	ReLU	Yes	0.5
Sixth Layer (Softmax Classifier)	**Dense**				**Activation**
	4				Softmax
Parameters					
Batch size	64				
No. of epochs	100				
Optimizer	Adam, SGD, RMSProp				
Initial learning	1e-2				

rate	
Total parameters	2,124,071
Trainable parameters	2,123,207
Non-trainable parameters	864

Table 1: Summary of layers of VADNet model

Three different types of optimizers—SGD, Adam and RMSProp are used to train the model. The sparse categorical cross entropy loss function is the one that is employed. Separable convolution 2D layer is favoured in the suggested deep learning design as it can lower the number of parameters without sacrificing speed. This is justified by the provision factorization, which reduces the number of parameters and, consequently, the size of the model and calculation time. In order to optimize, separable convolutional 2D (Eq. 3) takes advantage of the pointwise (Eq. 1) and depth wise (Eq. 2) convolution variations.

$$\mathrm{pc} = \sum_{\mathrm{m}}^{\mathrm{M}} \mathrm{W_m}.\mathrm{y}(\mathrm{i},\mathrm{j},\mathrm{m}) \quad (1)$$

$$\mathrm{dc}(\mathrm{W},\mathrm{y})_{(\mathrm{i},\mathrm{j})} = \sum_{\mathrm{k,l}}^{\mathrm{K,L}} \mathrm{W}_{(\mathrm{k,l})} \odot \mathrm{y}(\mathrm{i}+\mathrm{k},\mathrm{j}+\mathrm{l}) \quad (2)$$

$$\mathrm{sc}(\mathrm{W_p},\mathrm{W_d},\mathrm{y})_{(\mathrm{i},\mathrm{j})} = \mathrm{pc}_{(\mathrm{i},\mathrm{j})}(\mathrm{W_p},\mathrm{dc}_{(\mathrm{i},\mathrm{j})}(\mathrm{W_d},\mathrm{y})) \quad (3)$$

In this case, the standard convolutional method is the pointwise approach. A single kernel is used in the depth-wise variation, although additional parameters are produced. To guarantee spatial invariance and optimize feature maps, 2D layers with max pooling are employed. Because the pool size is (2,2), processing is done depending on the pooling window. It uses Eq. 4's expression of subsampling to optimize feature maps.

$$a_j = \tanh(\beta \sum_{N\,X\,N} a_i^{n\,xn} + b) \quad (4)$$

It takes into account multiplying inputs by a β-denoted trainable scalar. Next, trainable bias, indicated by the letter b, is added, and non-linearity is used to pass the result. In Equation 5, the max pooling function is shown.

$$a_j = \max_{N\,XN}(a_i^{n\,xn} u(n,n)) \quad (5)$$

The max pooling layer computes the maximum of the neighbourhood by utilizing a window function for a given input patch, which is represented as u(x,y). An optimized feature map is the end result. The normalizing method used in between the layers of the suggested VADNet model is called batch normalization. Multiple batches are utilized in place of whole data to facilitate learning, adapt to learning rates, and expedite the training phenomena. As in Eq. 6, the batch normalization is performed.

$$z^N = (\frac{z - m_z}{s_z}) \quad (6)$$

Neuronal output is represented by the symbols m_z for the mean and S_z for the standard deviation. Dense layers are also employed in the VADNet model that is suggested. Every neuron in the dense layer receives input from every other neuron in the layer above. It is known as a dense layer as a result. It can categorize a picture by using the convolutional layers' output that it receives.

3.4 Proposed Algorithm

We proposed an algorithm named Learning based Video Abnormality Detection (LbVAD) which exploits VADNet for efficient detection of video abnormalities. UCF-Crime is the benchmark dataset used for our empirical study.

Algorithm: Learning based Video Abnormality Detection (LbVAD)

Input: UCF-Crime dataset D

Output: Abnormality detection results R, performance statistics P

1. Begin
2. (T1, T2)←SplitData(D')
3. (T1, T2)←Augmentation(T1, T2)
4. Construct VADNet model
5. Add layer 1 (conv, BN, max pooling)
6. Add layer 2 (conv, BN, max pooling)
7. Add layer 3 (conv, BN, conv, BN, max pooling)
8. Add FC layer 1 (flatten, dense, ReLU, BN, dropout)
9. Add FC layer 2 (dense, ReLU, BN, dropout)
10. Add softmax classifier (dense, softmax)
11. model←TrainVADNet(T1)
12. R←TestVADNet(T2)
13. P←Evaluation(ground truth, R)
14. Display R
15. Display P
16. End

Algorithm 1: Learning based Video Abnormality Detection (LbVAD)

As presented in Algorithm 1, it takes UCF-Crime dataset [41] as input and follows a supervised learning method for detection and classification of abnormalities. It splits the dataset into 75% and 25% data towards training and testing respectively. The data is then subjected to analysis to know whether there is need for data augmentation. To improve quality of training data augmentation techniques such as flipping, shifting, zoom and rotation range are used. A deep learning model VADNet is built for training and performing abnormality detection. The model is used for training and then performing video abnormality detection.

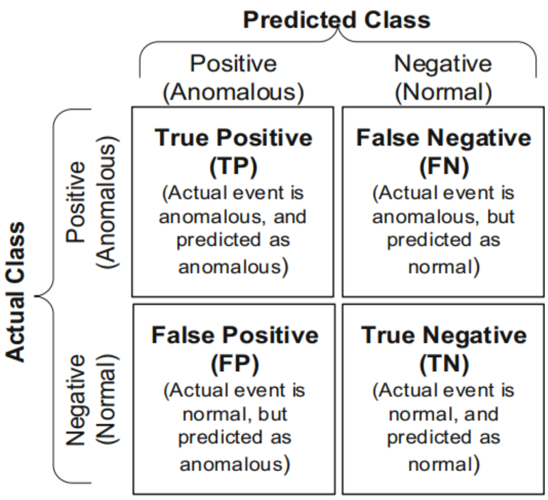
3.5 Dataset Details

Dataset named UCF-Crime [41] is used in our experiments. The dataset is rich with realistic incidents in videos. This dataset is widely used in video abnormality detection applications as in [42]. It has 13 classes of abnormalities. But in this research we used 4 classes for evaluating our proposed deep learning architecture VADNet.

3.6 Evaluation Method

We used accuracy as an important evaluation metric which is based on the cases presented in confusion matrix visualized in Figure 3.

Fig 3: Confusion matrix



Four cases such as TP, TN, FP and FN are presented in confusion matrix. They are used to assess performance of the proposed model in terms of accuracy metric expressed in Eq. 7.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

4. Experimental Results

Parameter	**Values set for Different Optimizers**		
	SGD	**Adam**	**RMSProp**
Batch Size	64	64	64
Learning Rate	le-2	le-0.001	le-0.001
Number of Epochs	100	100	100

Table 2: Parameters used in evaluation of VADNet with three different optimizers

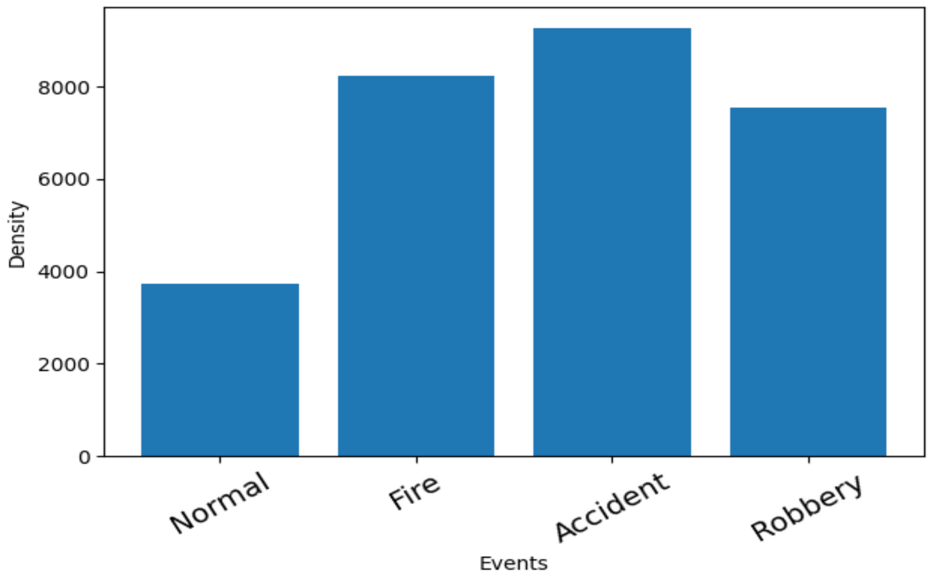


Fig 4: Data distribution in UCF-Crime dataset

5. Data Loading

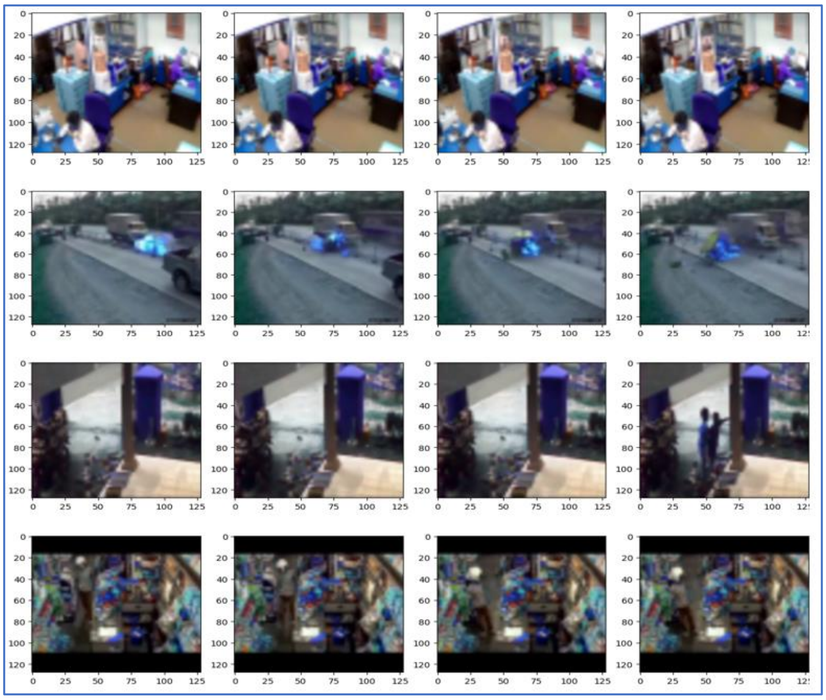


Fig 5: An excerpt from UCF-Crime dataset

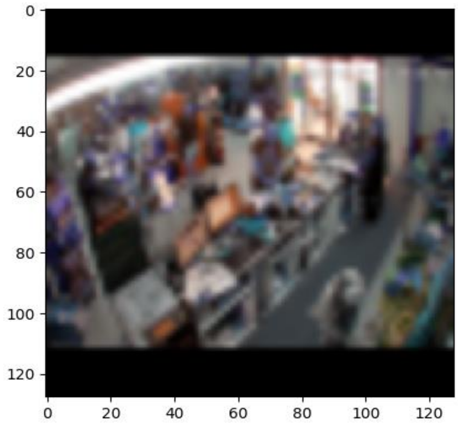


Fig 6: Given input video for analysis

Display each channel

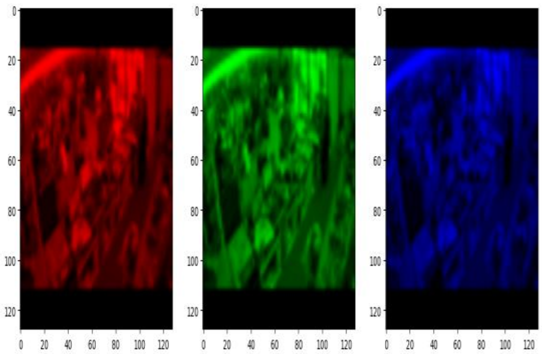


Fig 7: Different channels extracted from input video

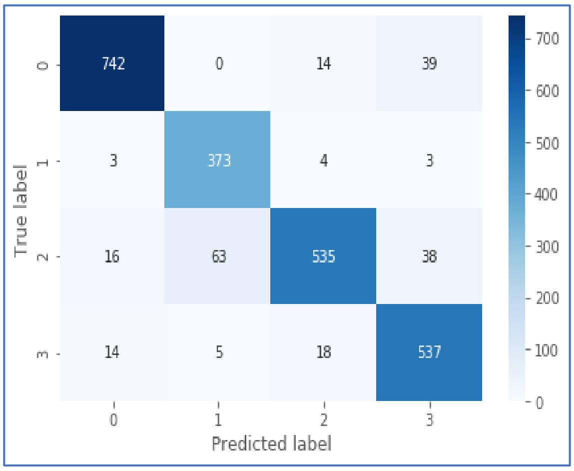


Fig 8: Confusion matrix showing performance of VADNet with adam optimizer

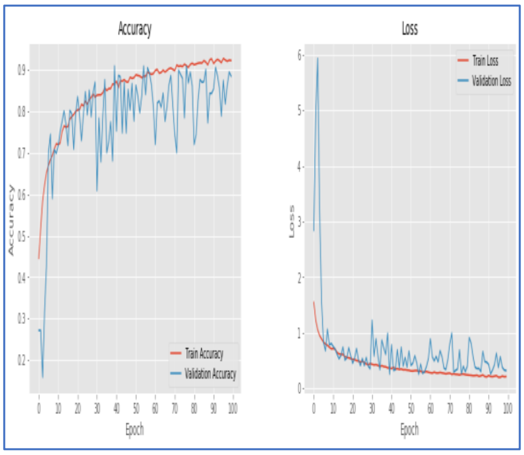


Fig 9: Performance of VADNet with adam optimizer

SGD

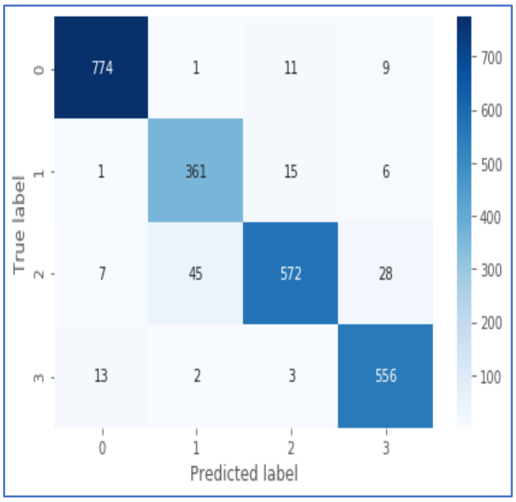


Fig 10: Confusion matrix showing performance of VADNet with SGD optimizer

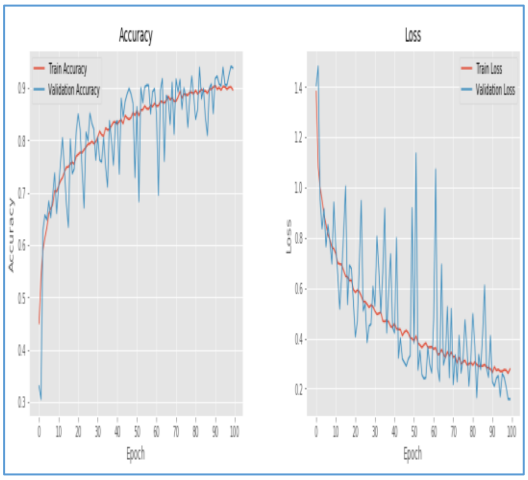


Fig 11: Performance of VADNet with SGD optimizer

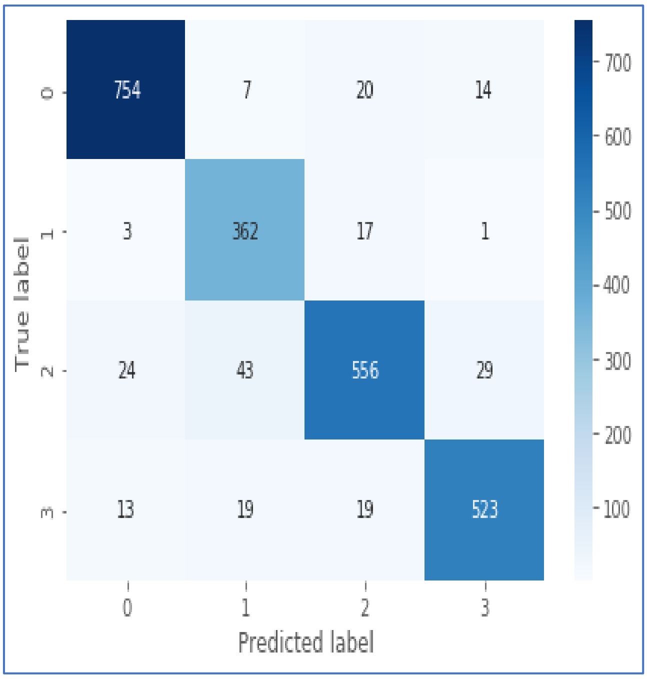


Fig 12: Confusion matrix showing performance of VADNet with RMSProp optimizer

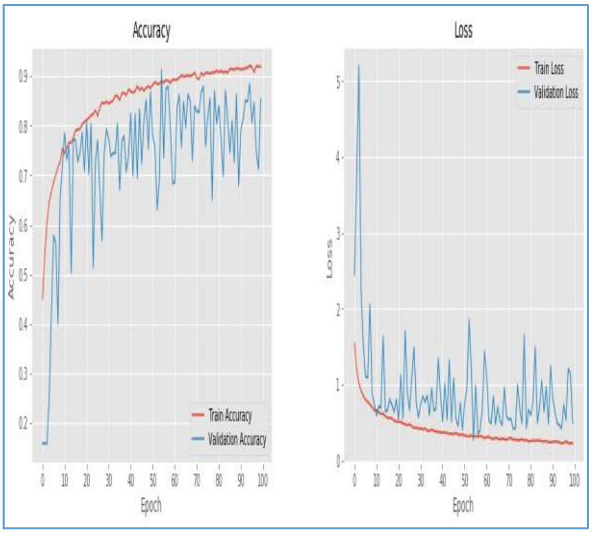


Fig 13: Performance of VADNet with RMSProp optimizer

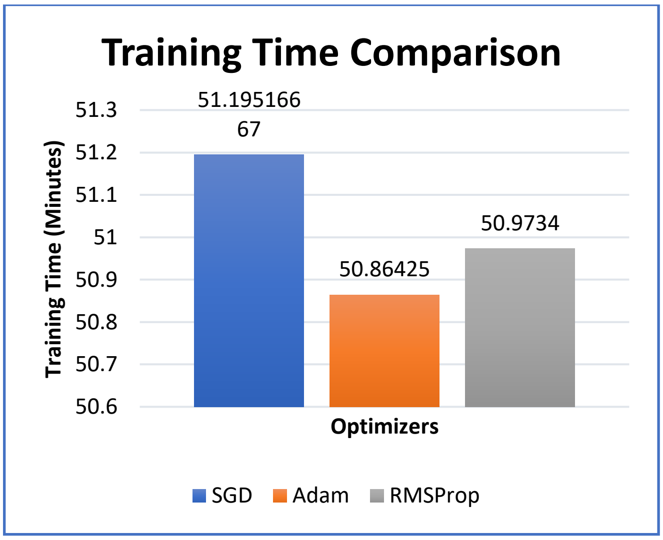


Fig 14: Training time comparison of VADNet with different optimizers

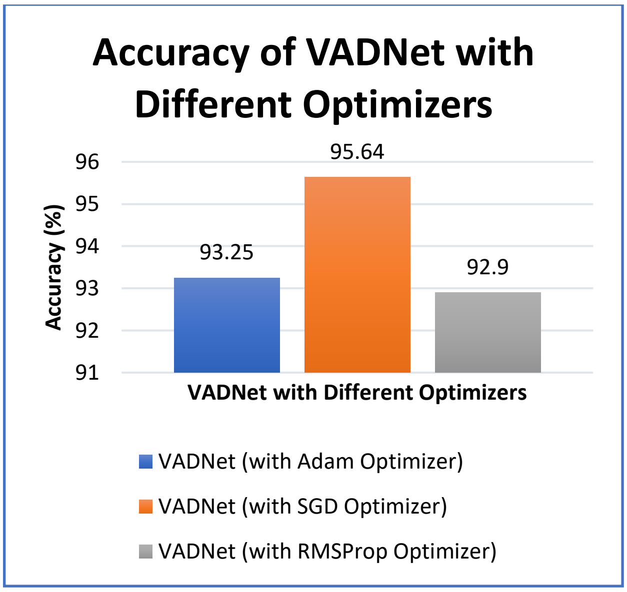


Fig 15: Accuracy of VADNet with different optimizers

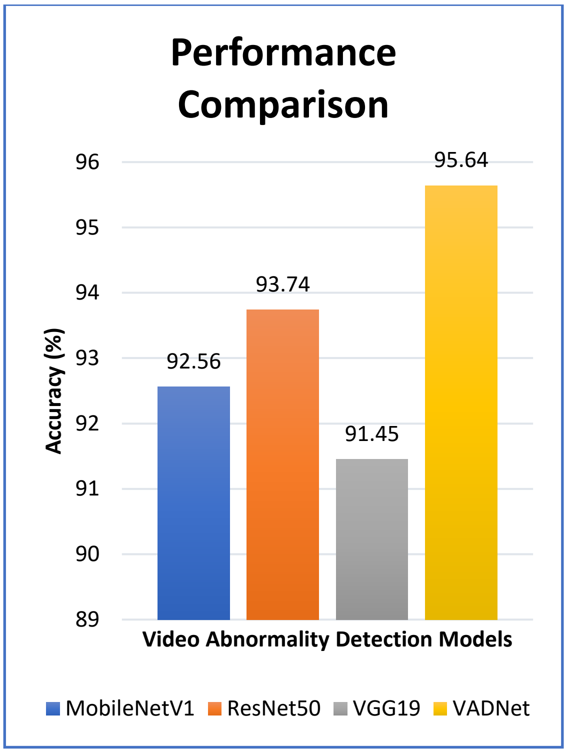


Figure 16: Performance comparison of VADNet with existing deep learning models

6. Conclusion and Future Work

In this paper, we proposed a novel deep learning architecture known as Video Abnormality Detection Net (VADNet) for detection of abnormalities from surveillance videos. VADNet is a CNN variant designed for leveraging detection performance. We proposed an algorithm named Learning based Video Abnormality Detection (LbVAD) which exploits VADNet for efficient detection of video abnormalities. This model has multiple layers and the fully connected layers followed by softmax activation function for abnormality classification. VADNet is evaluated with different optimizers such as adam, SGD and RMSProp. UCF-Crime is the benchmark dataset used for our empirical study. Our experimental results revealed that VADNet outperforms existing CNN variants like MobileNetV1, ResNet50 and VGG19 models with highest accuracy 95.64%. in future, we intend to propose deep learning models for crowd detection and crowd behaviour analysis from surveillance videos.

References

[1] Khaleghi Ali and Moin Mohammad Shahram (2018). Improved anomaly detection in surveillance videos based on a deep learning method, IEEE, pp.73–81. doi:10.1109/RIOS.2018.8406634

[2] Rashmika Nawaratne, Damminda Alahakoon and Daswin De Silvaand Xinghuo Yu. (2019). Spatiotemporal Anomaly Detection Using Deep Learning for Real-Time Video Surveillance. IEEE, pp.1-10.

[3] Kavikuil K. and Amudha J. (2018). Leveraging Deep Learning for Anomaly Detection in Video Surveillance. First International Conference on Artificial Intelligence and Cognitive Computing, pp.239–247. doi:10.1007/978-981-13-1580-0_23

[4] Khosro Rezaee, Sara Mohammad Rezakhani, Mohammad R. Khosravi and Mohammad Kazem Moghimi. (2021). A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance. Personal and Ubiquitous Computing, doi:10.1007/s00779-021-01586-5

[5] Nayak Rashmiranjan, Pati Umesh Chandra and Das Santos Kumar (2021). A comprehensive review on deep learning-based methods for video anomaly detection. Image and Vision Computing, 106, 104078–. doi:10.1016/j.imavis.2020.104078

[6] Sardar Waqar Khan, Qasim Hafeez, Muhammad Irfan Khalid and Roobaea Alroobа. (2022). Anomaly Detection in Traffic Surveillance Videos Using Deep Learning. MDPI, pp.1-28.

[7] John Gatara Munyua, Geoffrey Mariga Wambugu and Stephen Thiiru Njenga. (2021). A Survey of Deep Learning Solutions for Anomaly Detection in Surveillance Videos. International Journal of Computer and Information Technology. 10(5), pp.1-8.

[8] Aberkane Sabrina and Elarbi Mohamed (2019). Deep Reinforcement Learning for Real-world Anomaly Detection in Surveillance Videos, IEEE, pp.1–5. doi:10.1109/ISPA48434.2019.8966795

[9] Sabah Abdulazeez Jebur 1,2, Khalid A. Hussein 3 , Haider Kadhim Hoomod 3 , Laith. (2023). Review on Deep Learning Approaches for Anomaly Event Detection in Video Surveillance. MDPI, pp.1-22.

[10] Cosar, Serhan; Donatiello, Giuseppe; Bogorny, Vania; Garate, Carolina; Alvares, Luis O. and Bremond, Francois (2016). Towards Abnormal Trajectory and Event Detection in Video Surveillance. IEEE Transactions on Circuits and Systems for Video Technology, pp.1–14. doi:10.1109/TCSVT.2016.2589859

[11] Giannakeris, Panagiotis; Kaltsa, Vagia; Avgerinakis, Konstantinos; Briassouli, Alexia; Vrochidis, Stefanos; Kompatsiaris, Ioannis. (2018). Speed Estimation and Abnormality Detection from Surveillance Cameras. , IEEE, pp.93–936. doi:10.1109/CVPRW.2018.00020

A. Balasundaram, Golda Dilip, M. Manickam, Arun Kumar Sivaraman and K. Gu. (2022). Abnormality Identification in Video Surveillance System using DCT. Intelligent Automation & Soft Computing, pp.1-12.

[12] Ahmed, Sk. Arif; Dogra, Debi Prosad; Kar, Samarjit; Roy, Partha Pratim (2018). Surveillance scene representation and trajectory abnormality detection using aggregation of multiple concepts. Expert Systems with Applications, 101, pp.43–55. doi:10.1016/j.eswa.2018.02.013

[13] Lin, Weiyao; Zhang, Yihao; Lu, Jiwen; Zhou, Bing; Wang, Jinjun; Zhou, Yu (2015). Summarizing surveillance videos with local-patch-learning-based abnormality detection, blob sequence optimization, and type-based synopsis. Neurocomputing, 155, pp.84–98. doi:10.1016/j.neucom.2014.12.044

[14] Meng, B., Zhang, L., Jin, F., Yang, L., Cheng, H., & Wang, Q. (2017). Abnormal Events Detection Using Deep Networks for Video Surveillance. Cognitive Systems and Signal Processing, pp.197–204. doi:10.1007/978-981-10-5230-9_22

[15] Mujtaba Asad;He Jiang;Jie Yang;Enmei Tu;Aftab Ahmad Malik; (2021). Multi-Stream 3D latent feature clustering for abnormality detection in videos . Applied Intelligence, pp.1–18. doi:10.1007/s10489-021-02356-9

[16] Ye, Ou; Deng, Jun; Yu, Zhenhua; Liu, Tao; Dong, Lihong (2020). Abnormal Event Detection via Feature Expectation Subgraph Calibrating Classification in Video Surveillance Scenes. IEEE Access, 8, pp.97564–97575. doi:10.1109/access.2020.2997357

[17] Ko, Kwang-Eun; Sim, Kwee-Bo (2018). Deep convolutional framework for abnormal behaviour detection in a smart surveillance system. Engineering Applications of Artificial Intelligence, 67(), 226–234. doi:10.1016/j.engappai.2017.10.001

[18] Prawiro, Herman; Peng, Jian-Wei; Pan, Tse-Yu; Hu, Min-Chun (2020). Abnormal Event Detection in Surveillance Videos Using Two-Stream Decoder, IEEE, pp.1–6. doi:10.1109/ICMEW46912.2020.9105987

[19] Li, Ang; Miao, Zhenjiang; Cen, Yigang; Zhang, Xiao-Ping; Zhang, Linna; Chen, Shiming (2020). Abnormal event detection in surveillance videos based on low-rank and compact coefficient dictionary learning. Pattern Recognition, 108, pp.1–16. doi:10.1016/j.patcog.2020.107355

[20] Paul, Manoranjan; Haque, Shah M E; Chakraborty, Subrata (2013). Human detection in surveillance videos and its applications - a review. EURASIP Journal on Advances in Signal Processing, 2013(1), pp.1–16. doi:10.1186/1687-6180-2013-176

[21] Mabrouk, Amira Ben; Zagrouba, Ezzeddine (2017). Abnormal behaviour recognition for intelligent video surveillance systems: a review. Expert Systems with Applications, pp.1–55. doi:10.1016/j.eswa.2017.09.029

[22] Yan Shiyang, Smith Jeremy S., Lu Wenjin and Zhang Bailing. (2018). Abnormal Event Detection from Videos using a Two-stream Recurrent Variational Autoencoder. IEEE Transactions on Cognitive and Developmental Systems, pp.1–14. doi:10.1109/TCDS.2018.2883368

[23] Wang, Siqi; Zeng, Yijie; Liu, Qiang; Zhu, Chengzhang; Zhu, En; Yin, Jianping (2018). [ACM Press 2018 ACM Multimedia Conference - Seoul, Republic of Korea (2018.10.22-2018.10.26)] 2018 ACM Multimedia Conference on Multimedia Conference - MM '18 - Detecting Abnormality without Knowing Normality, pp.636–644. doi:10.1145/3240508.3240615

[24] Chong, Y. S., and Tay, Y. H. (2017). Abnormal Event Detection in Videos Using Spatiotemporal Autoencoder. Lecture Notes in Computer Science, pp.189–196. doi:10.1007/978-3-319-59081-3_23

[25] Tong Li;Xinyue Chen;Fushun Zhu;Zhengyu Zhang;Hua Yan; (2021). Two-stream deep spatial-temporal auto-encoder for surveillance video abnormal event detection. Neurocomputing, pp.1–15. doi:10.1016/j.neucom.2021.01.097

[26] Leyva, Roberto; Sanchez, Victor; Li, Chang-Tsun (2017). The LV dataset: A realistic surveillance video dataset for abnormal event detection, IEEE, pp.1–6. doi:10.1109/IWBF.2017.7935096

[27] Zhou, Fuqiang; Wang, Lin; Li, Zuoxin; Zuo, Wangxia; Tan, Haishu (2019). Unsupervised Learning Approach for Abnormal Event Detection in Surveillance Video by Hybrid Autoencoder. Neural Processing Letters, pp.1–15. doi:10.1007/s11063-019-10113-w

[28] Cui, Lili; Li, Kehuang; Chen, Jiapin; Li, Zhenbo (2011). Abnormal event detection in traffic video surveillance based on local features, IEEE, pp.362–366. doi:10.1109/CISP.2011.6099933

[29] Dong, Nan; Jia, Zhen; Shao, Jie; Xiong, Ziyou; Li, Zhipeng; Liu, Fuqiang; Zhao, Jianwei; Peng, PeiYuan (2010). Traffic Abnormality Detection through Directional Motion Behavior Map, IEEE, pp.80–84. doi:10.1109/avss.2010.61

[30] Leyva, Roberto; Sanchez, Victor; Li, Chang-Tsun (2017). Abnormal event detection in videos using binary features, IEEE, pp.621–625. doi:10.1109/TSP.2017.8076061

[31] Varadarajan, Jagannadan; Odobez, Jean-Marc (2009). Topic models for scene analysis and abnormality detection, IEEE, pp.1338–1345. doi:10.1109/iccvw.2009.5457456

[32] Leyva, Roberto; Sanchez, Victor; Li, Chang-Tsun (2018). Fast Detection of Abnormal Events in Videos with Binary Features, IEEE, pp.1318–1322. doi:10.1109/ICASSP.2018.8461759

[33] Sharma, M. K.; Sheet, D.; Biswas, P. K. (2020). Spatiotemporal deep networks for detecting abnormality in videos. Multimedia Tools and Applications, pp.1–32. doi:10.1007/s11042-020-08786-w

[34] Dr. Rajesh Sharma and Dr. Akey Sungheetha. (2021). An Efficient Dimension Reduction based Fusion of CNN and SVM Model for Detection of Abnormal Incident in Video Surveilla. Journal of Soft Computing Paradigm (JSCP). 3(2), pp.55-69.

[35] Ravanbakhsh, Mahdyar; Nabi, Moin; Sangineto, Enver; Marcenaro, Lucio; Regazzoni, Carlo; Sebe, Nicu (2017). Abnormal event detection in videos using generative adversarial nets, pp.1577–1581. doi:10.1109/ICIP.2017.8296547

[36] Yihao Zhang, ; Weiyao Lin, ; Guangwei Zhang, ; Chuanfei Luo, ; Dong Jiang, ; Chunlian Yao, (2014). A new approach for extracting and summarizing abnormal activities in surveillance videos, IEEE, pp.1–5. doi:10.1109/ICMEW.2014.6890537

[37] A.Niranjil Kumar and Dr. C.Sureshkumar. (2014). Abnormal Crowd Detection and Tracking in Surveillance Video Sequences. International Journal of Advanced Research in Computer and Communication Engineering. 3(9), pp.1-5.

[38] Popoola, O. P.; Kejun Wang, (2012). Video-Based Abnormal Human Behavior Recognition A Review. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6), pp.865–878. doi:10.1109/tsmcc.2011.2178594

[39] Abdulamir A. Karim and Narjis Mezaal Shati. (2017). Abnormality Detection using K-means Data Stream Clustering Algorithm in Intelligent Surveillance System. Journal of AL-Qadisiyah for computer science and mathematics. 9(1), pp.82 – 98.

[40] UCF-Crime Dataset. Retrieved from https://www.crcv.ucf.edu/projects/real-world/

[41] 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