Grenze International Journal of Engineering and Technology, June Issue

Analysis of Deep Models for Abnormal Event Detection in Video Surveillance

Susheel Kumar Gadeda Goudar[1], Indira Priya Darshini[2] and V N Manjunath Aradhya[3]
[1]DoS in Computer Science, University of Mysore, Mysuru 570006
[2,3]Dept. of Computer Applications, JSS Science and Technology University, Mysuru, 570006
Email: aradhya@jssstuniv.in

***Abstract*—The happening of abnormal events compared to normal events is hardly rare. When abnormal events occur in a place that is under a surveillance system, those surveillance systems are capable of detecting the type of abnormal events by differentiating them from normal events, which will enhance the security system and be useful in rapid incident response. Scene classification from surveillance video is intrinsically complex, because of different dynamics in nature and human intervention. In this paper to address these challenges, the proposed method combines advanced deep learning techniques to effectively process and analyse video data. Convolutional Neural Networks (CNNs) such as VGG19 and ResNet-101 are employed to capture spatial features, while Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units and Gated Recurrent Units (GRUs) are utilized to model temporal dependencies within the video sequences. Additionally, Support Vector Machines (SVMs) are explored for their classification capabilities in this context. Experiments were conducted on a comprehensive dataset comprising CCTV footage representative of each event category. The proposed system shows significant promise in distinguishing between different normal and different kinds of abnormal events.**

***Keywords*—abnormal events, surveillance system, deep learning, CCTV footage.**

I. Introduction

Nowadays video surveillance has become an integral part of security in every sector. Initially, these systems were purely analogue, relying on closed-circuit television (CCTV) to monitor specific areas. These systems required manual observation, where security personnel had to continuously watch live video feeds to detect suspicious activities. While effective to an extent, this manual approach was time-consuming, error-prone, and limited by human capabilities, particularly when faced with monitoring large areas or multiple feeds simultaneously.

The digital revolution paved the way for change in the use of video surveillance where the analogue equipment was replaced by high-definition cameras, networked video feeds and centralized storage systems. However, the sudden increase of video data acceleratively changed the problem of analysis and approaches to tackle large video data into a new direction. Human operators could not react on time due to the overload of data, meaning that human operators often failed to respond quickly and missed key events due to overload, especially in conditions that require the analysis of large amounts of information.

Grenze ID: 01.GIJET.11.2.757
© Grenze Scientific Society, 2025

Recent advancements in artificial intelligence (AI) have introduced significant innovations in the field of video surveillance. AI-driven surveillance systems now have the ability to automatically analyse video footage and predict whether the footage is normal or abnormal. This automation will dramatically reduce the dependency on human operators. Even though it detects between normal and abnormal events it lags in classifying the type of abnormal event that is going on. Relatively less research was done on the classification of the type of abnormal events.

For the above-mentioned limitation to come up with a solution, the proposed system aims to develop a robust and efficient system capable of detecting and classifying more accurately of normal and type of abnormal events within the footage of surveillance.

Deep Nural Networks will lay as the best architectures for these difficult classification tasks such as the complete variation in the background of similar events in different places, the dynamics of nature vary, the position of the occurrence of the events and the uncertainty of the human behaviour in those events. Deep Learning Models are used because of their capability for automatic feature extraction, hierarchical feature representation, joint optimization and state-of-the-art performance.

The proposed system utilizes three datasets to train the models for classifying the different types of events in surveillance videos. Visual Geometry Group – 19(VGG-19) and Residual Neural Network – 101(ResNet-101), two deep convolutional neural networks have been explored for feature extraction purposes. Support Vector Machine (SVM) - a supervised machine learning algorithm, Long Short Term Memory (LSTM) – a specialized type of RNN and Gated recurrent units (GRUs) – a variant of RNNs are explored for the purpose of classification.

The organization of the paper goes like: section II comprises an overview of existing methods and research relevant to abnormal event classification. Section III details the methodology of the proposed system.

II. Literature Review

A Violence Detection Model Using an Attention-Based CNN-LSTM Structure is proposed by Soheil Vosta and Kin-Choong Yow [2]. This study presents KianNet, a unique novel model that combines the ResNet50 and ConvLSTM architectures with a multi-head self-attention layer to detect violent incidents inside recorded events. Robust feature extraction is made possible by the use of ResNet50, and exploiting the temporal dependencies in the video sequences is made simpler by ConvLSTM. Additionally, the multi-head self-attention layer improves the model's discriminatory skill and ability to concentrate on pertinent spatiotemporal regions. On the UCF-Crime dataset, the model's accuracy in binary classification is 97.48% whereas, on the RWF dataset, it is 96.21%.

Deep Learning Approach for Suspicious Activity Detection from Surveillance Video is addressed in [3]. This work puts forward a system that will use the footage obtained from CCTV cameras for monitoring students' activities in a campus and send messages to the corresponding authority when any suspicious event occurs. Here for the model training, the model uses the videos from the CAVIAR dataset, KTH dataset, YouTube videos and videos taken from campus. As part of pre-processing, frames are extracted from the captured videos and then trained with pre-trained model VGG-16 is used for feature extraction and then LSTM architecture is used for classification. The system achieves 87.15% accuracy with the trained model while it predicts the frames as suspicious (mobile phone using inside the campus, fighting or fainting) or normal (walking, running) class.

Abnormal Event Detection in Public Places by Deep Learning Methods is described in [4]. This work presents the implementation and comparison of three different methods: convolutional long short-term memory (CLSTM) autoencoders, convolutional autoencoders (CA), and one-class support vector machines (SVM) to detect the abnormal over the dataset UCSD Ped1. One-class SVM achieves the highest accuracy of 65.82%, CLSTM autoencoder 62.75%, and Convolutional Autoencoder 61.85% accuracy.

An Integration of Handcrafted Features for Violent Event Detection in Videos is proposed in [5]. This paper outlines the work in the direction of detection of violent and non-violent events for which they have used the Hockey Fight and Violent-Flow dataset. For experimentation, the GHOG (Global Histograms of Oriented Gradients), HOFO (Histogram of Optical Flow Orientation), and GIST feature descriptors are used and have integrated these hand-crafted feature descriptors. The work uses SVM classifier for the categorization of violent and non-violent events. The model achieves 96.93% accuracy for Hockey Fight dataset and 95.45 accuracy for the Violent-Flow dataset.

Anomaly recognition from surveillance videos using 3D convolution neural network proposes a method that utilizes deep learning-based approaches, particularly focusing on 3D ConvNets [6] to address the challenge of anomalous activity recognition from surveillance videos. The suggested method focuses on multi-class anomaly identification, in contrast to previous deep learning models that frequently handled just binary classification and older handcrafted methods that struggled with complicated event detection. By using 3D ConvNets to extract spatiotemporal characteristics and spatial augmentation, the model is better able to capture temporal and spatial contexts. Experiments on the UCF-Crime dataset show that this approach significantly improves performance, with an 82% accuracy.

Violent Video Event Detection Based on Integrated LBP and GLCM Texture Features outlines the work on the task of detection of violent and nonviolent events for which they have used the Hockey Fight and Violent-Flow dataset [7]. This work proposes the method based on the texture feature descriptors. For experimentation LBP, GLCM and features fusion descriptors are used for the detection of violent events. Five supervised Classifiers, Decision Tree (DT), Discriminant Analysis (DA), Logistic Regression (LR), SVM and K-Nearest Neighbor (KNN) are used for the categorization of violent and nonviolent events. The results show that SVM consistently outperforms other classifiers across both HF and VF datasets, achieving the highest accuracy of 91.51% and AUC of 93.60% on the HF dataset with fusion features, while LR also performs well, particularly on the VF dataset.

Abnormal event detection for video surveillance using deep one-class learning' [2017] is described in [8] This work proposes a Deep One-Class (DOC) model that integrates CNN with one-class SVM for abnormal event detection in video surveillance. The CNN learns high-dimensional representations of normal events, while the one-class SVM layer acts as both a discriminator and an optimization objective. A robust loss function derived from the one-class SVM optimizes the model in an end-to-end manner, simplifying the process and achieving a global optimal solution. Experimental results on a public dataset show the DOC model's superior performance compared to state-of-the-art methods, demonstrating its effectiveness in detecting abnormal events.

Abnormal Event Detection in Videos Based on Deep Neural Networks employs a Variational Autoencoder (VAE) which enforces the hidden layer representation of the normal samples to follow by a Gaussian distribution [9]. The method proposes training the VAE in such a way that this distribution accurately represents normal samples. For the purpose of abnormality detection, the VAE encodes test samples onto a hidden layer, the representation of test sample is afterwards generated and tested for its likeness of conforming to the Gaussian distribution. A sample is said to be abnormal when its representation is found to be above a particular detection threshold. This method obtains a frame-level auccracy of 92.3% for the UCSD dataset and 82.1% for the Avenue dataset.

Automatic Traffic Accident Detection System Using ResNet and SVM is described by Aman Kumar Agarwal, Kadamb Agarwal et al. [10]. In order to identify possible accidents from traffic surveillance cameras, this work suggests a model that combines clustering and classification techniques. It does this by first dividing all videos into smaller shots based on scene changes, then extracting important frames from each shot using the histogram differences of consecutive frames, and finally calculating the distance between vehicles. A ResNet50 architecture is used to extract features from these crucial frames, and then K Means clustering is used to create a Bag of Visual Words (BOVW). This is then fed into a Support Vector Machine (SVM) classifier to determine if it is an accident in the video. The accuracy achieved by the model is 94.14%. The experiments were carried out with a dataset of 32 videos, 16 of which had accidents and 16 of which did not.

III. Proposed Methodology

A. Methodology

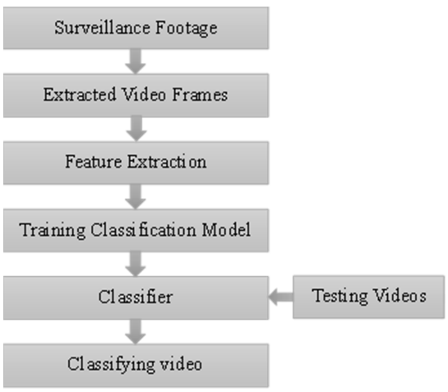


Figure 1. General architecture of video surveillance system

In video classification, considering only a frame of the event for classification purposes is insufficient, as the circumstances demand the entire sequence to be captured for the accurate interpretation of the events. Therefore, implementing a video surveillance system using deep learning needs a multi-stage approach, starting with the data collection and ending with the successful classification of abnormal events. The general architecture for the video surveillance is given in figure 1. This architecture outlines the process of classification of abnormal events in surveillance videos. The system begins by taking raw surveillance footage from which individual frames are extracted and these frames are fed to the feature extraction phase, where important information is gathered from each frame to capture key characteristics of the scene. For this purpose, two feature extraction techniques namely VGG19 and ResNet-101 have been used. The extracted features are fed into a neural model for training a classifier. The model learns patterns from the data that help in distinguish between 4 different types of events: Accident, Explosion, Fighting and Normal. In this work, three different classifiers namely SVM, LSTM and GRUs have been used for classification. After training, the model is tested on new videos, where the same process is repeated. The classifier analyses the frames and determines to which class the event belongs. The output of the system is the detection of specific events within the tested videos.

B. Dataset Description

In the development of the proposed system, two datasets are employed along with the surveillance footage videos downloaded from YouTube to train and test the models for classifying the different types of events in surveillance videos. UCF-Crime Dataset: This formed our primary dataset, providing a comprehensive collection of CCTV footage across four classes—accident, explosion, fighting, and normal. RWF-2000 Dataset: This dataset provides videos specifically of fighting and normal events, providing additional data for these classes. YouTube videos: The downloaded videos are focused on the accident and explosion classes, used to improve the dataset of these classes. Combining all datasets together provides the ability to detect these specific anomalies.

C. Model Architecture

For the classification of abnormal events, we incorporated two different feature extraction models VGG19 and ResNet-101 to extract spatial features. SVM, LSTM and GRUs are taken for the purpose of building the classifiers.

Visual Geometry Group-19 (VGG-19)

VGG-19 was for feature extraction because of its ability to capture detailed spatial features through its deep convolutional layers. Its architecture is simple, which is based on small 3*3 filters and is significantly good at identifying fine details in video frames. The availability of pre-trained VGG19 models on large datasets like ImageNet provides a solid basis for extracting key features efficiently.

Residual Neural Network-101 (ResNet-101)

For feature extraction, ResNet-101 was employed because of its deep 101-layer architecture. It is good at capturing complex spatial features from video frames. Because of its residual connections, ResNet-101 can be trained on deeper networks without concern of vanishing gradient problems during training. This enables it to extract complex attributes from the video frames.

Support Vector Machine (SVM)

SVM was taken for classification purposes since it was performed accurately during binary and multi-class classification tasks.

SVM is designed to detect the optimal hyperplane for members of different classes and therefore, suits well in the classification of normal versus abnormal events on the features extracted by VGG19 and ResNet-101. It also performs well on high-dimensional data, which is essential when classifying the rich features of these extraction models.

Long Short-Term Memory (LSTM)

LSTM was selected for classification because it is designed for sequential data like video frames, which will capture long-term dependencies. As video analysis involves the comprehension of the sequence of events, it was reasonable to utilize LSTMs in this case. By identifying how events unfold over a series of frames, its memory cells let it retain crucial information from the collected characteristics, improving classification accuracy.

Gated Recurrent Units (GRUs)

GRUs were selected for classification as they represent a less resource-demanding option as compared to LSTMs while still being able to model temporal dependencies in sequential data. One option that GRUs offer is improved training efficiency and reduced resource requirements because they will use gating mechanisms to control the flow of information and subsequently classify video events from the extracted features.

D. Proposed Models

We proposed the combination of nine different models by varying the classification models with each feature extraction method. Among nine models three models are of the ensemble learning method, which combines the features extracted from both the feature extraction model and trained the individual classification models.

IV. Result and Analysis

A. Dataset Preparation

To carry out the experiment with altogether of three datasets, the final usable dataset has been prepared. The preparation steps were carefully carried out to ensure the taking of only the relevance of the data. First, the manual trimming of the videos from the datasets is done to take only the specific portions that contain the abnormal events. Once the videos were trimmed, they were converted into individual frames with a resolution of 224*224 pixels. The frames were extracted at a rate of 10 frames per second (fps) for the accident, explosion, and fighting classes, and 5 fps for the normal class. Every frame was stored in their respective class folder. The model training and evaluation were conducted for two types of training and testing split percentage. The videos were manually split into training and testing sets in the ratio of 75:25 and 80:20 is shown from Table 1 to Table 20 respectively.

B. Experiment Results

The proposed models were experimented on the described datasets. Testing was conducted using two different dataset split configurations: 75:25 and 80:20, ensuring a comprehensive assessment of model performance under varying training-to-testing data ratios. The results obtained from these experiments are presented for comparative analysis and validation.

Table I: Results with VGG19_SVM

	Precision	Recall	F1-Score
Accident	0.79	0.76	0.78
Explosion	0.80	0.78	0.79
Fighting	0.75	0.70	0.73
Normal	0.58	0.65	0.61

Table II: Results with VGG19_LSTM

	Precision	Recall	F1-Score
Accident	0.83	0.79	0.81
Explosion	0.91	0.84	0.87
Fighting	0.82	0.72	0.77
Normal	0.62	0.76	0.68

Table III: Results with VGG19_GRUs

	Precision	Recall	F1-Score
Accident	0.83	0.77	0.80
Explosion	0.93	0.85	0.89
Fighting	0.77	0.73	0.75
Normal	0.60	0.71	0.65

Table IV: Results with ResNet_SVM

	Precision	Recall	F1-Score
Accident	0.83	0.82	0.83
Explosion	0.88	0.83	0.86
Fighting	0.84	0.77	0.80
Normal	0.69	0.78	0.73

TABLE V: RESULTS WITH RESNET_LSTM

	Precision	Recall	F1-Score
Accident	0.82	0.88	0.85
Explosion	0.94	0.79	0.86
Fighting	0.88	0.76	0.81
Normal	0.66	0.80	0.72

TABLE VI: RESULTS WITH RESNET_GRUS

	Precision	Recall	F1-Score
Accident	0.82	0.94	0.87
Explosion	0.96	0.77	0.85
Fighting	0.92	0.78	0.84
Normal	0.71	0.87	0.78

TABLE VII: RESULTS WITH VGG19 AND RESNET _SVM

	Precision	Recall	F1-Score
Accident	0.85	0.79	0.82
Explosion	0.86	0.85	0.85
Fighting	0.82	0.72	0.77
Normal	0.62	0.76	0.68

TABLE VIII: RESULTS WITH VGG19 AND RESNET _LSTM

	Precision	Recall	F1-Score
Accident	0.82	0.91	0.86
Explosion	0.76	0.84	0.80
Fighting	0.72	0.94	0.82
Normal	0.84	0.41	0.55

TABLE IX: RESULTS WITH VGG19 AND RESNET _GRUS

	Precision	Recall	F1-Score
Accident	0.83	0.94	0.88
Explosion	0.90	0.84	0.87
Fighting	0.79	0.92	0.85
Normal	0.79	0.60	0.68

The overall accuracy of the models is presented in the table 10.

TABLE X: ACCURACY OF THE PROPOSED MODELS OF 75:25 SPLIT

Proposed Models	**Result**
VGG19_SVM	71.31%
VGG19_LSTM	76.78%
VGG19_GRUs	75.38%
ResNet-101_SVM	71.46%
ResNet-101_LSTM	80.08%
ResNet-101_GRUs	82.43%
VGG19 & ResNet-101_SVM	76.69%
VGG19 & ResNet-101_LSTM	76.67%
VGG19 & ResNet-101_GRUs	73.69%

TABLE XI: RESULTS WITH VGG19_SVM

	Precision	Recall	F1-Score
Accident	0.96	0.75	0.84
Explosion	0.82	0.80	0.81
Fighting	0.84	0.68	0.76
Normal	0.66	0.91	0.76

TABLE XII: RESULTS WITH VGG19_LSTM

	Precision	Recall	F1-Score
Accident	0.91	0.79	0.84
Explosion	0.95	0.77	0.85
Fighting	0.93	0.72	0.81
Normal	0.66	0.99	0.79

TABLE XIII: RESULTS WITH VGG19_GRUS

	Precision	Recall	F1-Score
Accident	0.96	0.70	0.81
Explosion	0.86	0.81	0.83
Fighting	0.84	0.79	0.81
Normal	0.70	0.90	0.79

TABLE XIV: RESULTS WITH RESNET_SVM

	Precision	Recall	F1-Score
Accident	0.90	0.81	0.85
Explosion	0.90	0.80	0.85
Fighting	0.84	0.68	0.75
Normal	0.64	0.87	0.73

TABLE XV: RESULTS WITH RESNET_LSTM

	Precision	Recall	F1-Score
Accident	0.90	0.90	0.90
Explosion	0.92	0.77	0.84
Fighting	0.92	0.74	0.82
Normal	0.66	0.93	0.79

TABLE XVI: RESULTS WITH RESNET_GRUS

	Precision	Recall	F1-Score
Accident	0.85	0.95	0.90
Explosion	0.95	0.79	0.86
Fighting	0.80	0.85	0.83
Normal	0.76	0.73	0.74

TABLE XVII: RESULTS WITH VGG19 AND RESNET _SVM

	Precision	Recall	F1-Score
Accident	0.97	0.77	0.85
Explosion	0.91	0.82	0.86
Fighting	0.91	0.67	0.77
Normal	0.63	0.96	0.76

TABLE XVIII: RESULTS WITH VGG19 AND RESNET _LSTM

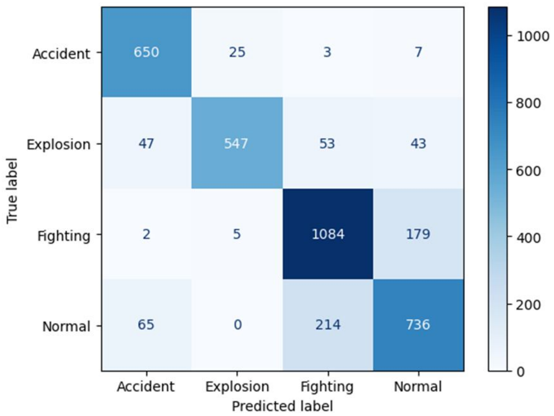
	Precision	Recall	F1-Score
Accident	0.82	0.92	0.86
Explosion	0.72	0.82	0.77
Fighting	0.80	0.99	0.89
Normal	0.88	0.48	0.62

TABLE XIX: RESULTS WITH VGG19 AND RESNET _GRUS

	Precision	Recall	F1-Score
Accident	0.80	0.95	0.87
Explosion	0.71	0.82	0.76
Fighting	0.81	0.96	0.88
Normal	0.95	0.53	0.68

The overall accuracy of the models is presented in table 20.

TABLE XX: ACCURACY OF THE PROPOSED MODELS OF 80:20 SPLIT



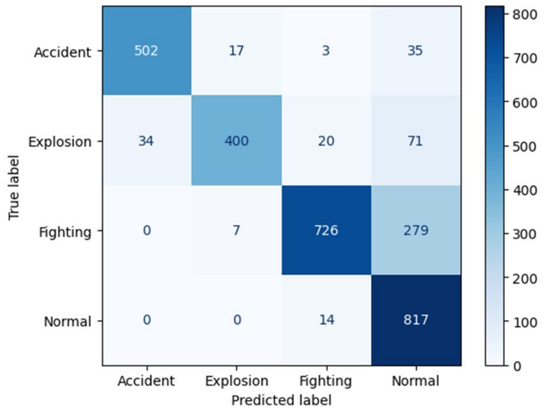


Figure 2: Confusion matrix of ResNet101_GRUs model of 75:25 split Figure 3: Confusion matrix of ResNet101_GRUs model of 80:20 split

C. Result Analysis

Overall, the results show consistent trends across both dataset splits, with GRUs outperforming other classifiers in most cases. For the 75:25 split, ResNet-101_GRUs achieves the highest accuracy at 82.43%, closely followed by its 80:20 split counterpart at 83.59%, highlighting GRUs' superior ability to capture temporal dependencies. LSTMs also deliver strong performance, with ResNet-101_LSTM achieving 80.08% for the 75:25 split and 82.70% for the 80:20 split.

SVM-based models consistently underperform, as seen with ResNet-101_SVM scoring 71.46% (75:25) and 78.05% (80:20), reflecting their limitations in temporal learning tasks. Hybrid models combining VGG19 and ResNet-101 generally improve performance, with VGG19 & ResNet-101_GRUs scoring 73.69% (75:25) and 80.96% (80:20), although they do not outperform standalone ResNet-101 models. The results indicate higher accuracies with an 80:20 split, likely due to a larger training set improving model generalisation. Overall, the ResNet-101_GRUs model emerges as the most effective, demonstrating robust performance across both splits. Confusion matrix for the splits is shown in Figure 2 and 3.

V. CONCLUSION

Abnormal event classification from the surveillance system will make the proper use of the system in time rather than only using footage after manually knowing the incidents. The results of the proposed systems demonstrate the efficiency of deep learning techniques in the complex task of abnormal event classification from video surveillance data. The proposed models have shown promising outcomes in classifying abnormal events that enhance the reliability of the surveillance system.

In particular, the feature extraction through ResNet-101 followed by classification using an advanced recurrent neural network namely GRUs, provided a nuanced understanding of temporal dynamics in the video sequences, leading to improved classification accuracy of 82.43% on the 75:25 split and 83.59% on the 80:20 split. The results are futuristic and will enable further research in this area even better with emphasis on the detection of abnormal events.

Several areas could be explored further to enhance and expand its capabilities such as the incorporation of real-time processing, increasing the dataset for more robustness, and expanding to additional classes for increasing versatility, and multimodal data integration for additional support for most accuracy.

REFERENCES

[1] Dr. Anala M.R, Malika Makker, and Aakanksha Ashok “Anomaly Detection in Surveillance Videos” 2019 26th International Conference on High Performance Computing, Data, and Analytics (HiPC) Workshop (HiPCW).

[2] Soheil Vosta and Kin-Choong Yow “KianNet: A Violence Detection Model Using an Attention-Based CNN-LSTM Structure” in IEEE Digital Object Identifier 10.1109/ACCESS.2023.3339379, 4 December 2023.

[3] Amrutha C.V, C. Jyotsna, and Amudha J "Deep Learning Approach for Suspicious Activity Detection from Surveillance Video“Second International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020).

[4] Mattaparti Satya Bhargavi, J. V. Bibal Benifa, and Rishav Jaiswal "Abnormal Event Detection in Public Places by Deep Learning Methods“published in the Advances in Intelligent Systems and Computing series, volume 1335. (2021)

[5] B.H. Lohithashva, V.N. Manjunath Aradhya, and D.S. Guru “An Integration of Handcrafted Features for Violent Event Detection in Videos” was published in the 2021 edition of Recent Trends in Computational Intelligence Enabled Research, spanning pages 295-305.

[6] Ramna Maqsood, Usama Ijaz Bajwa, Gulshan Saleem, Rana Hammad Raza and Muhammad Waqas Anwar “Anomaly recognition from surveillance videos using 3D convolution neural network”. Multimedia Tools and Applications (2021) 80:18693–18716

[7] B.H. Lohithashva1, V.N. Manjunath Aradhya1 and D.S. Guru “Violent Video Event Detection Based on Integrated LBP and GLCM Texture Features”. Revue d'Intelligence Artificielle Vol. 34, No. 2, April, 2020, pp. 179-187

[8] Jiayu Sun. Jie Shao. Chengkun He “Abnormal event detection for video surveillance using deep one-class learning”.(2017) Multimed Tools Appl https://doi.org/10.1007/s11042-017 5244-2

[9] Qinmin Ma “Abnormal Event Detection in Videos Based on Deep Neural Networks”. Hindawi Scientific Programming Volume 2021, Article ID 6412608, 8 pages. https://doi.org/10.1155/2021/6412608

[10] Aman Kumar Agrawal, Kadamb Agarwal, Jitendra Choudhary, Aradhita Bhattacharya, Srihitha Tangudu, Nishkarsh Makhija and Rajitha B “Automatic Traffic Accident Detection System Using ResNet and SVM”. 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)

[11] Lohithashva, B.H., Aradhya, V.N.M., Basavaraju, H.T., Harish, B.S. (2019). “Unusual crowd event detection: An approach using probabilistic neural network”. In Information Systems Design and Intelligent Applications, pp. 533-542.

[12] Ovhal, K. B., Patange, S. S., Shinde, R. S., Tarange, V. K., & Kotkar, V. A. (2017). “Analysis of Anomaly Detection Techniques in Video Surveillance”. In Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017) (pp. [specific pages]). IEEE Xplore. ISBN: 978-1-5386-1959-9.

[13] Lei, J., Sun, W., Fang, Y., Ye, N., Yang, S., & Wu, J. (2024). “A Model for Detecting Abnormal Elevator Passenger Behaviour Based on Video Classification”. Electronics, 13(2472). https://doi.org/10.3390/electronics13132472.

[14] Dadula, C. P., & Dadios, E. P. (2015). “Neural Network Classification for Detecting Abnormal Events in a Public Transport Vehicle”. In Proceedings of the 8th IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. [specific pages]). IEEE. ISBN: 978 1-5090-0360-0.

[15] Jebur, S. A., Alzubaidi, L., Hussein, K. A., Hoomod, H. K., Saihood, A. A., & Gu, Y. (2024). “A Scalable and Generalized Deep Learning Framework for Anomaly Detection in Surveillance Videos”. arXiv preprint arXiv:2408.00792.