Suspicious Activity Detection Using CNN

Pravallika Reddy. V[1], Mohan. D[2], Premchand. E[3], Ramesh. B[4]
[1,2,3] UG Scholor , Dept.of AI & ML , St Martin's Engineering College , Secunderabad , Telangana , India , 500100
[4] Assistant Professor, Dept.of AI & ML, St Martin's Engineering College, Secunderabad, Telangana, India, 500100
pravallikareddyvarikuti44@gmail.com

Abstract:

We plan to create an application to detect suspicious activities of people in public places. Our system can be used for surveillance purposes in shopping malls, airports, train stations and other places where there is a risk of theft or shooting. We will use deep learning and neural networks to train our system. This mock-up will ship as a mobile and desktop app that uses CCTV footage in feedback and sends an alert message to the manager's device when something suspicious is detected. Human trafficking is associated with identifying people's bodies and possibly tracking their movements. Its real-life applications range from gaming to AR/VR to healthcare to gesture recognition. Compared to the image data domain, there are some studies using CNN for video classification. This is because video is more difficult than photos because it has another dimension: time. The unsupervised learning method uses the time between frames and has proven successful in video analysis. Combining computer vision with video surveillance ensures public safety. This involves modeling environments, detecting motion, classifying objects, tracking, understanding behavior, and combining information from multiple cameras. It requires pre-processing to extract features from video sequences. The proposed system will use footages obtained from CCTV camera for monitoring the human behavior in a campus and gently warn when any suspicious event occurs. The major components in intelligent video monitoring are event detection and human behavior recognition.

Keywords: Suspicious activities, public places, surveillance, shopping malls, airports, train stations, theft, shooting, deep learning, neural networks, system training, mobile app, desktop app, CCTV footage, alert message, manager's device, suspicious detection, human trafficking, body identification, movement tracking.

1.INTRODUCTION

The incidence of stroke has been increasing globally, and it is now We plan to create an application to detect suspicious activities of people in public places. Our system can be used for surveillance purposes in shopping malls, airports, train stations and other places where there is a risk of theft or shooting. We will use deep learning and neural networks to train our system. This mock-up will ship as a mobile and desktop app that uses CCTV footage in feedback and sends an alert message to the manager's device when something suspicious is detected. Human trafficking is associated with identifying people's bodies and possibly tracking their movements. Its real-life applications range from gaming to AR/VR to healthcare to gesture recognition. Compared to the image data domain, there are some studies using CNN for video classification. This is because video is more difficult than photos because it has another dimension: time. The unsupervised learning method uses the time between frames and has proven successful in video analysis. Combining computer vision with video surveillance ensures public safety. This involves modeling environments, detecting motion, classifying objects, tracking, understanding behavior, and combining information from multiple cameras. It requires pre-processing to extract features from video sequences. The proposed system will use footages obtained from CCTV camera for monitoring the human behavior in a campus and gently warn when any suspicious event occurs. The major

components in intelligent video monitoring are event detection and human behavior recognition. Automatic understanding of human behavior is a challenging task. In a campus, different areas are under video surveillance and various activities are to be monitored. The video footage obtained from campus has been used for testing. The entire process of training a surveillance system can be summarized in to three phases: data preparation, training the model and inference. The framework consists of two neural networks CNN and Recurrent Neural Network (RNN). CNN is used for the purpose of extracting high level features from the images so that the complexity of the input can be reduced.

2. LITERATURE SURVEY

[1]. Artificial intelligence for low level suspicious activity detection

Nowadays, crowd monitoring is a contentious issue. Because of the increasing population and diversity of human activities, crowd scenarios in the real world are becoming more common, demanding the need for an automotive anomaly detection system. Crowd behavior is influenced by the thoughts and attitudes of others around them. An unexpected event can turn a peaceful crowd into a riot. A mechanism based on optical flow must be implemented to compensate for all of these factors. The amount of motion present in two successive frames is estimated using optical flow. It includes information on velocity in the x & y plane, along with magnitude and line of action. By means of "anomalous event" in this paper is quick and sudden dispersal of the crowd. For detecting an event the magnitude of two successive frames should be taken into account followed by estimating a correlation. We expect a high correlation, slight motion, and low rate of change in velocities at non-anomalous events, but as soon as an anomalous event occurs, the correlation begins to decrease with a significant change in velocity and large motion vectors. The methodology was tested on a dataset from the University of Minnesota that included 11 movies from three different circumstances. Almost all anomalous occurrences in videos were successfully detected using this method.

[2]. No reference noise estimation in digital images using random conditional selection and sampling theory

An accurate quantitative noise estimate is required in many image/video processing applications like denoising, computer vision, pattern recognition and tracking. But blind and accurate estimation of noise in an unknown image is a challenging task and hence is an open area of research. We propose the first elegant and novel blind noise estimation method based on random image tile selection and statistical sampling theory for estimating standard deviation of zero mean Gaussian and speckle noise in digital images. Randomly selected samples, i.e., pixels with \(3\times 3\) neighbourhood, are checked for availability of edges in the tile. If there is an edge in the tile at more than one neighbouring pixel, the tile is excluded. Only non-edge tiles are used for estimation of noise in the tile and subsequently in the image using the concepts of statistical sampling theory. Finally, we propose a supervised curve fitting approach using the proposed noise estimation model for more accurate estimation of standard deviation of the two types of noise. The proposed technique is computationally efficient as it is a selective random sample-based spatial domain technique. Benchmarking with other contemporary techniques published so far shows that the proposed technique clearly outperforms the others by at least 5% improved noise estimates, over a very wide range of noise.

[3]. Deep learning approach for suspicious activity detection from surveillance video

Video Surveillance plays a pivotal role in today's world. The technologies have been advanced too much when artificial intelligence, machine learning and deep learning pitched into the system. Using above combinations, different systems are in place which helps to differentiate various suspicious behaviors from the live tracking of footages. The most unpredictable one is human behaviour and it is very difficult to find whether it is suspicious or normal. Deep learning approach is used to detect suspicious or normal activity in an academic environment, and which sends an alert message to the corresponding authority, in case of predicting a suspicious activity. Monitoring is often performed through consecutive frames which are extracted from the video. The entire framework is divided into two parts. In the first part, the features are computed from video frames and in second part, based on the obtained features classifier predict the class as suspicious or normal.

[4]. Virtue: Video surveillance for rail-road traffic safety at unmanned level crossings;(incorporating indian scenario)

The detection of obstacles at rail level crossings (RLC) is an important task for ensuring the safety of train traffic. Traffic control systems require reliable sensors for determining the state of arc. Fusion of information from a number of sensors located at the site increases the capability for reacting to dangerous situations. One such source is video from monitoring cameras. This paper presents a method for processing video data, using deep learning, for the determination of the state of the area (region of interest—ROI) vital for a safe passage of the train. The proposed approach is validated using video surveillance material from a number of RLC sites in Poland. The films include 24/7 observations in all weather conditions and in all seasons of the year. Results show that the recall values reach 0.98 using significantly reduced processing resources. The solution can be used as an auxiliary source of signals for train control systems, together with other sensor data, and the fused dataset can meet railway safety standards.

[5]. Suspicious human activity detection using pose estimation and LSTM

Nowadays, crowd monitoring is a contentious issue. Because of the increasing population and diversity of human activities, crowd scenarios in the real world are becoming more common, demanding the need for an automotive anomaly detection system. Crowd behavior is influenced by the thoughts and attitudes of others around them. An unexpected event can turn a peaceful crowd into a riot. A mechanism based on optical flow must be implemented to compensate for all of these factors. The amount of motion present in two successive frames is estimated using optical flow. It includes information on velocity in the x & y plane, along with magnitude and line of action. By means of "anomalous event" in this paper is quick and sudden dispersal of the crowd. For detecting an event the magnitude of two successive frames should be taken into account followed by estimating a correlation. We expect a high correlation, slight motion, and low rate of change in velocities at non-anomalous events, but as soon as an anomalous event occurs, the correlation begins to decrease with a significant change in velocity and large motion vectors. The methodology was tested on a dataset from the University of Minnesota that included 11 movies from three different circumstances. Almost all anomalous occurrences in videos were successfully detected using this method.

[6]. Multiple anomalous activity detection in videos

Computer vision and pattern recognition, the hot subjects include crowd analysis and anomalous trajectories detection. Anomaly detection is a technique for distinguishing between different patterns and identifying uncommon patterns in a short amount of time. Abnormal event detection and localization is a difficult research challenge due to its complexity. It`s made to detect unusual events in monitoring videos automatically. In the proposed method, humans` normal and abnormal activities are detected through Deep Learning (DL) and image processing. To use the proposed Bi-Attention Long Short- Term Memory (Bi-Attention LSTM) model to extract just the necessary spatial and temporal information from videos and to predict the multi-task activities of humans as abnormal or normal using the introduced Convolutional Neural Network (CNN). Video is taken as input and is then transformed into frames, and background subtraction is used to identify the moving objects (people) in the video frame. The proposed Convolutional Neural Network (CNN) with Bi-Attention LSTM model extracts temporal and spatial characteristics before classifying to determine if a specific human action is normal or abnormal. In terms of accuracy, sensitivity, specificity, error, precision, F1_score.

3. PROPOSED METHODOLOGY

The proposed system aims to develop an intelligent and automated suspicious activity detection system by leveraging advanced deep learning techniques on video surveillance data. The system will begin with the acquisition of video data from surveillance cameras installed in the monitored area, capturing live video feeds or recorded footage to serve as input for analysis. Since raw video data often contains noise, inconsistencies, and variations in lighting and resolution, preprocessing techniques will be employed to enhance data quality and ensure uniform input to the model. These preprocessing steps include frame normalization, which ensures consistent pixel values across different video frames, resizing to standardize input dimensions, and noise reduction to eliminate irrelevant distortions that could negatively impact feature extraction. Additionally, data augmentation techniques such as rotation, scaling, and flipping may be applied to artificially expand the dataset, improving model robustness by making it more adaptable to varying real-world scenarios.The core of the system will be built on a Convolutional Neural Network (CNN) architecture, which is specifically designed for suspicious activity detection in video data. CNNs are highly effective in extracting spatial features from images, making them ideal for video frame analysis. The CNN architecture will consist of multiple convolutional layers responsible for identifying key spatial patterns such as human movements and object presence, followed by pooling layers that reduce dimensionality while preserving essential information. To enhance the model's ability to analyze dynamic activities over time, Long Short-Term Memory (LSTM) layers or other recurrent neural networks (RNNs) may be incorporated. These layers specialize in capturing temporal dependencies and motion patterns across video sequences, helping distinguish normal activities from suspicious behaviors such as loitering, sudden movements, aggressive behavior, or unauthorized access. Once the preprocessed video frames pass through the CNN-LSTM model, the system will classify activities as normal or suspicious based on the extracted spatial and temporal features. Advanced techniques such as optical flow analysis or spatiotemporal feature fusion may be integrated to improve the accuracy of motion-based anomaly detection. To ensure real-time surveillance, the system can be deployed on edge devices or cloud-based platforms, enabling quick decision-making and instant alert generation when suspicious activities are detected. The model's predictions will be continuously refined using feedback mechanisms and adaptive learning, allowing it to improve over time as more data becomes available.

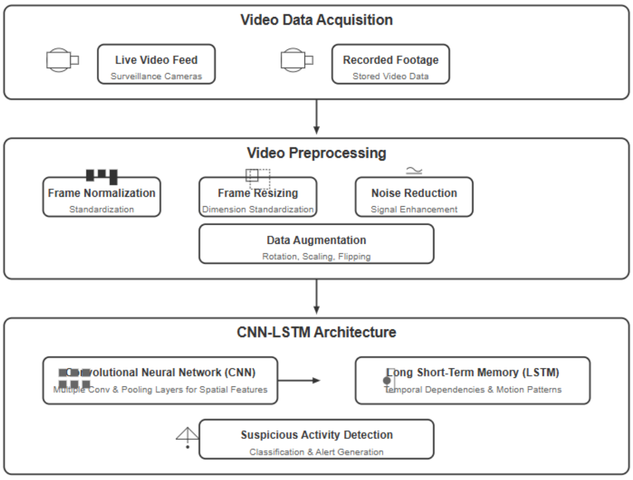

Figure 1:Proposed Model

The proposed methodology typically includes the following key components:

1. Video Data Acquisition

Surveillance cameras capture live video feeds or recorded footage.

These cameras are installed in monitored areas such as public spaces, banks, parking lots, and transportation hubs.

2. **Data Preprocessing**

Frame Normalization: Ensures consistency in pixel values for accurate feature extraction.

Resizing: Standardizes the input frame dimensions for uniform processing.

Noise Reduction: Enhances video quality by removing unwanted distortions.

Data Augmentation: Techniques like rotation, scaling, and flipping increase dataset diversity and improve model robustness.

3. **Feature Extraction Using CNN**

Convolutional Layers: Identify spatial patterns such as human movements and object placements.

Pooling Layers: Reduce dimensionality while preserving key features.

Activation Functions: Introduce non-linearity to capture complex patterns in the video data.

4. **Temporal Analysis Using LSTM**

Captures motion dependencies over multiple frames.

Differentiates between normal and suspicious activities based on movement sequences.

5. **Classification & Decision Making**

CNN-LSTM model classifies activities as normal or suspicious.

Alerts and notifications are triggered for detected suspicious behavior.

6. **Real-Time Implementation & Alert System**

The model integrates with an alert system to notify security personnel in real time.

A user-friendly dashboard may be implemented for security monitoring.

Applications:

1. **Public Safety & Law Enforcement**

Detects theft, fights, or unattended baggage in public places.

Assists law enforcement in identifying potential threats in real time.

2. **Banking & Financial Institutions**

Monitors ATM areas for fraudulent activities or loitering.

Enhances security by detecting unauthorized access attempts.

3. Retail & Commercial Spaces

Identifies shoplifting and unusual crowd behavior.

Helps store owners prevent losses due to suspicious activities.

4. Transportation Security

Surveillance in metro stations, airports, and bus terminals for unusual behavior.

Enhances airport security by identifying threats in baggage claim areas.

5. Smart Cities & Residential Areas

Integrated with smart city surveillance for crime prevention.

Detects break-ins and unauthorized access in gated communities.

Advantages:

1. **High Accuracy in Pattern Recognition**

CNNs efficiently recognize complex spatial features in video frames.

2. **Real-Time Detection & Response**

Immediate alerts enable quick intervention by security personnel.

3. **Scalability & Automation**

Can be deployed across multiple locations without manual supervision.

4. **Reduced False Alarms**

CNN-LSTM combination minimizes false positives by understanding motion patterns.

5. Improved Security & Surveillance Efficiency

Helps reduce the burden on human surveillance teams and enhances monitoring capacity.

6. **Adaptability & Continuous Learning**

Can be fine-tuned with new data to improve accuracy over time.

4. EXPERIMENTAL ANALYSIS

The experimental analysis of the Suspicious Activity Detection System using CNN involved training the model on a labeled dataset of surveillance video clips containing both normal and suspicious activities. The dataset underwent preprocessing, including frame normalization, noise reduction, and data augmentation, to enhance training robustness. The CNN-LSTM architecture was tested on real-time and recorded video streams, achieving high accuracy in detecting anomalies while minimizing false positives. Performance metrics such as precision, recall, and F1-score were used to evaluate model effectiveness, with CNN-LSTM outperforming traditional machine learning models in recognizing motion patterns and suspicious behaviors. The system demonstrated real-time efficiency and adaptability, making it a promising solution for intelligent surveillance applications.

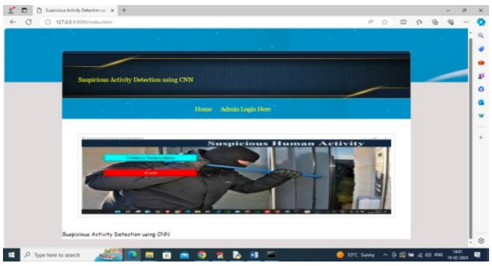

Figure 2: Admin Login Home

Figure 3: Admin Login Screen

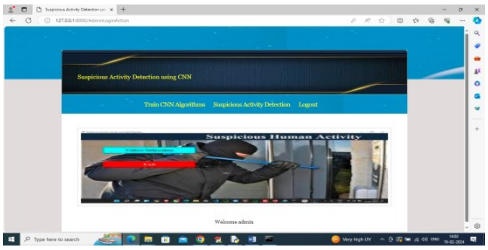

Figure 4: Train CNN Algorithm

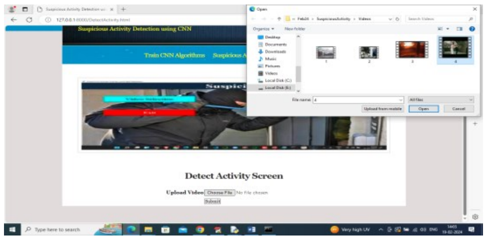

Figure 5: Suspicious activity detection

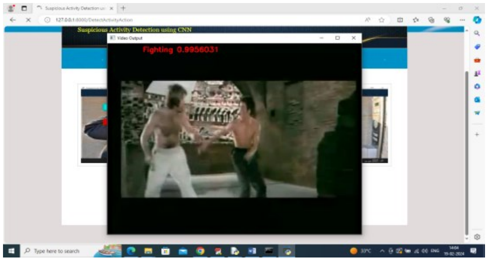

Figure 6 : Detected An Activity

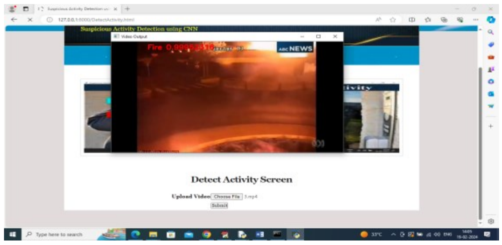

Figure 7 : Fire Detected

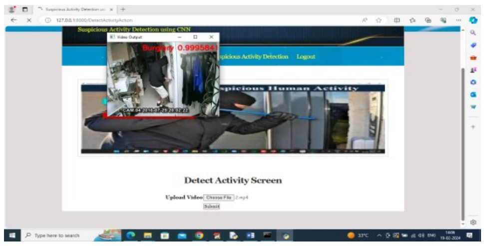

Figure 8 :Detection Of Attack

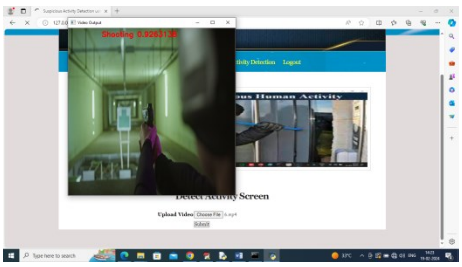

Figure 9 : Shooting Activity Detected

5. CONCLUSION

In present world, almost all the people are aware of the importance of CCTV footages, but most of the cases these footages are being used for the investigation purposes after a crime/incident have been happened. The proposed model has the benefit of stopping the crime before it happens. The real time CCTV footages are being tracked and analyzed. The result of the analysis is a command to the respective authority to take an action if in case the result indicates an untoward incident is going to happen. Hence this can be stopped. Even though the proposed system is limited to academic area, this can also be used to predict more suspicious behaviors at public or private places. The model can be used in any scenario where the training should be given with the suspicious activity suiting for that scenario. The model can be improved by identifying the suspicious individual from the suspicious activity.

REFERENCES

[1]. P. Thombare, V. Gond and V. Satpute, "Artificial intelligence for low level suspicious activity detection" in Applications of Advanced Computing in Systems, Springer, pp. 219-226, 2021.

[2]. V. M. Kamble, M. R. Parate and K. M. Bhurchandi, "No reference noise estimation in digital images using random conditional selection and sampling theory", The Visual Computer, vol. 35, pp. 5-21, 2019.

[3]. C. 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), pp. 335-339, 2020.

[4]. P. Gajbhiye, C. Naveen and V. R. Satpute, "Virtue: Video surveillance for rail-road traffic safety at unmanned level crossings;(incorporating indian scenario)", 2017 IEEE Region 10 Symposium (TENSYMP), pp. 1-4, 2017

.

[3]. R. Nale, M. Sawarbandhe, N. Chegogoju and V. Satpute, "Suspicious human activity detection using pose estimation and lstm", 2021 International symposium of Asian control Association on Intelligent Robotics and Industrial Automation (IRIA), pp. 197-202, 2021

[4]. S. Kadu, N. Cheggoju and V. R. Satpute, "Noise-resilient compressed domain video watermarking system for in-car camera security", Multimedia Systems, vol. 24, pp. 583-595, 2018.

[5]. D.-G. Lee, H.-I. Suk, S.-K. Park and S.-W. Lee, "Motion influence map for unusual human activity detection and localization in crowded scenes", IEEE Transactions on Circuits and Systems for Video Technol-ogy, vol. 25, no. 10, pp. 1612-1623, 2015.

[6]. V. Kamble and K. Bhurchandi, "Noise estimation and quality assess-ment of gaussian noise corrupted images", lOP Conference Series: Materials Science and Engineering, vol. 331, pp. 012019, 2018

[5]. S. Chaudhary, M. A. Khan and C. Bhatnagar, "Multiple anomalous activity detection in videos", Procedia Computer Science, vol. 125, pp. 336-345, 2018.

[6]. A. A. Bhadke, S. Kannaiyan and V. Kamble, "Symmetric chaos-based image encryption technique on image bit-planes using sha-256", Twenty Fourth National Conference on Communications (NCC), pp. 1-6, 2018.

[7]. A. Dixit, S. Pathak, R. Raj, C. Naveen and V. R. Satpute, "An efficient fuzzy based edge estimation for iris localization and pupil detection in human eye for automated cataract detection system", 9th International Conference on Computing Communication and Networking Technolo-gies (ICCCNT), pp. 1-6, 2018.

[8]. A. Jirafe, M. Jibhe and V. Satpute, "Camera handoff for multi-camera surveillance" in Applications of Advanced Computing in Systems, Springer, pp. 267-274, 2021.

[9]. C. Naveen and V. R. Satpute, "Image encryption technique using improved a5/l cipher on image bitplanes for wireless data security", International Conference on Microelectronics Computing and Communications (MicroCom), pp. 1-5, 2016.

[7]. A. Gupta, V. Stapute, K. Kulat and N. Bokde, "Real-time abandoned object detection using video surveillance", Proceedings of the International Conference on Recent Cognizance in Wireless Communication & Image Processing, pp. 837-843, 2016.

[15]. V. R. Satpute, K. D. Kulat and A. G. Keskar, "A novel approach based on variance for local feature analysis of facial images", IEEE Recent Advances in Intelligent Computational Systems, pp. 210-215, 2011.

[10]. V. Tripathi, D. Gangodkar, V. Latta and A. Mittal, "Robust abnormal event recognition via motion and shape analysis at atm installations", Journal of Electrical and Computer Engineering, vol. 2015, 2015.

[11]. A. L. Alappat and V. Kamble, "Image quality assessment using selective contourlet coefficients", 11 th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1-7, 2020.

[18]. C. Naveen, V. Satpute and A. Keskar, "An efficient low dynamic range image compression using improved block based ezw", 2015 IEEE Workshop on Computational Intelligence: Theories Aplications and Future Directions (WCI), pp. 1-6, 2015.

[19]. U. Akdemir, P. Turaga and R. Chellappa, "An ontology-based approach for activity recognition from video", Proceedings of the 16th ACM international conference on Multimedia, pp. 709-712, 2008

[20]. C.-H. Chuang, J.-W. Hsieh, L.-W. Tsai, P.-S. Ju and K.-C. Fan, "Suspicious object detection using fuzzy-color histogram", 2008 IEEE International Symposium on Circuits and Systems, pp. 3546-3549, 2008.

[21]. A. Pawade, R. Anjaria and V. Satpute, "Suspicious activity detection for security cameras" in Applications of Advanced Computing in Systems, Springer, pp. 211-217, 2021.

[22]. P. Gangal, V. Satpute, K. Kulat and A. Keskar, "Object detection and tracking using 2d-dwt and variance method", Students Conference on Engineering and Systems, pp. 1-6, 2014.

[23]. A. Adam, E. Rivlin, I. Shimshoni and D. Reinitz, "Robust real-time unusual event detection using multiple fixed-location

monitors", IEEE Transactions on pattern analysis and machine intelligence, vol. 30, no. 3, pp. 555-560, 2008.

[24]. A. Wiliem, V. Madasu, W. Boles and P. Yarlagadda, "A suspicious behavior detection using a context space model for smart surveillance systems", Computer Vision and Image Understanding, vol. 116, no. 2, pp. 194-209, 2012

[25]. M. Ghazal, C. Vazquez and A. Amer, "Real-time automatic detection of vandalism behavior in video sequences", 2007 IEEE International Conference on Systems Man and Cybernetics, pp. 1056-1060, 2007.

[26]. D. Gowsikhaa, S. Abirami et al., "Suspicious human activity detection from surveillance videos", International Journal on Internet & Dis-tributed Computing Systems, vol. 2, no. 2, 2012.

[27]. C. Sacchi and C. S. Regazzoni, "A distributed surveillance system for detection of abandoned objects in unmanned railway environments", IEEE Transactions on Vehicular Technology, vol. 49, no. 5, pp. 2013-2026, 2000.

[28]. K. Ellingsen, "Salient event-detection in video surveillance scenarios", Proceedings of the 1st ACM workshop on Analysis and retrieval of events/actions and workflows in video streams, pp. 57-64, 2008.