International Journal of Gender, Science and Technology UGC Care Group I Journal
..ISSN: 2040-0748 Vol-13 Issue-01 May 2024

HUMAN ABNORMAL EVENT DETECTION FOR VEDIO SURVEILLANCE AND NON-SURVEILLANCE USING RNN

Harshitha Shree S[#1], Meenakshi R[#2], Monisha M[#3] Dr.MadhurikkhaS[#4]
[#1,2,3] Department Of Computer Science And Engineering,
[#4] Associate Professor / AI&DS
Jaya Sakthi Engineering College,Chennai,602024,Tamil Nadu,India
harshithashree2212@gmail.com,meenakshiramdass2003@gmail.com,monishamurugan1702@gmail.com, madhurikkha@gmail.com

Abstract*—With a special emphasis on the MobileNet architecture, Human Abnormal Event identification investigates the effectiveness of Convolutional Neural Network (CNN) models in abnormal event identification. Finding abnormal events is an important task in many fields, such as anomaly and surveillance detection, comparing MobileNet's performance with other well-known CNN models, like GoogleNet, VGG-16, and AlexNet. The precision with which anomalous occurrences in visual data can be identified by the MobileNet model is evaluated. We examine its advantages and disadvantages in this area via thorough testing and comparison with alternative architectures. A number of indicators are included in the assessment, including as robustness to various anomalies, accuracy, and computing efficiency. Our results show that MobileNet performs promisingly in tasks involving the detection of aberrant events. It is appropriate for real-time applications because it offers advantages in computing efficiency and displays accuracy levels that are competitive with those of current CNN architectures. We also show the domains in which MobileNet performs better or worse than other models, offering insights into its applicability for particular use cases. Overall, abnormal event detection advances knowledge of CNN-based methods for this type of detection and gives practitioners and projectors insightful information about which models are best for their needs. The results show how successful MobileNet may be and point to future directions for research to improve CNN models' performance in crucial domains.

***Keywords*—Surveillance, Deep Learning, Intelligent System, MobileNet ,vgg-16, GoogleNet ,Convolutional Neural Network, Recurrent Neural Networks, Object Detection, Optical Flow.**

I. INTRODUCTION

The observation of behavior, activities, or other changing information, usually pertaining to people, with the intention of regulating, protecting, influencing, or guiding them, is known as surveillance. Governments utilize surveillance for a variety of purposes, including intelligence collection, crime prevention, process, person, group, or object protection and crime scene examination. Suspicious activity detection is essential for a video surveillance system to be effective. Using such technologies to identify irregularities in human behavior has received more weight recently since it can offer hints while averting security breaches.

A person's features and behavioural patterns are crucial in identifying them. One important source of information for these identifications is visual data. Such visual data is provided via surveillance videos, which can be replayed for later use or seen in real time. Video analytics is not exempt from the effects of the recent "automation" movement. Applications for video analytics are numerous and include vehicle counting, people counting in crowded areas, motion detection, person identification, anomalous activity recognition, and human activity prediction The two elements in this area that are utilized to determine an individual's identity are referred to as face recognition and gait recognition, respectively. The more flexible of these two methods for automated human identification from surveillance footage is face recognition. Face recognition technology can be used to forecast a person's head orientation, which helps forecast their behavior. In numerous applications, such as identity verification, motion detection with facial recognition, and presence/absence detection at a certain location and time, motion recognition combined with facial recognition is highly beneficial.

Reconstruction or future prediction models are the foundation of yet another growing concept in the Abnormal Event Detection (AED) framework Auto-encoders trained by reconstruction-based algorithms are typically trained on normal data, where problematic frames in video sequences are identified by their larger reconstruction errors in comparison to normal frames. Large prediction errors are used in inference to identify abnormal frames since they are infrequent events that have distinct probability distributions from the training data. Other prediction-based AED algorithms train a network to anticipate the next frame based on the previous frame using raw RGB data or optical flow features.

The strength of these algorithms resides in their ability to make use of the spatial and temporal appearance and motion information. This allows them to identify a wide range of abnormal events, including strange behaviors like fighting, chasing, and jumping as well as abnormal object classifications like cars, bicycles, and skateboards. Additionally, there is no requirement for labeling because the training data is typical, which drastically lowers labor expenses. These algorithms, however, are difficult to train properly and insensitive to tiny aberrant classes.

This work suggests an integrated AED framework, as illustrated in Fig. 1, that combines several criteria to identify various kinds of aberrant events based on videos. Two parallel processing branches make up our framework. Pedestrians and other objects are classified in the first stream using the object detector, which extracts class information. The body pose data is then extracted using a pose estimation technique. A graph convolutional neural network processes the reduced pose information into high-level features following its first fusion with class information.

The initial integrated pose and class-based stream has the potential to enhance the precision of pose-level detection outcomes and effectively identify anomalous events unrelated to human activity. By using a flownet, the second stream retrieves the optical flow findings. Next, U-net functions as a generator, processing both the raw RGB data and the optical flow features. Moreover, the notion of adversarial learning uses a discriminator to forecast if the subsequent frame is authentic or fraudulent. With its ability to enhance look and motion description and make up for the shortcomings of the first stream, the second stream can improve abnormal event identification in situations involving small objects and busy spaces. To determine if the tested data are aberrant or normal, a final fusion of their normality scores from the two streams would be performed.

We show that RGB-based networks may be effectively enhanced for prediction and classification by a number of features. However, posture data in motion and appearance features can be compensated for using raw RGB data and derived optical flow findings. The primary contributions of this paper are:

1.Our proposal is a comprehensive framework for prediction fusion and classification that can identify different kinds of anomalous occurrences in security footage.

2.Even in busy settings, our AED algorithm exhibits enhanced sensitivity towards anomalous occurrences due to the use of motion features.

3.Our in-depth tests show how reliable and efficient the suggested framework is in AED.

4.We examine the many kinds of anomalous occurrences found in multiple open AED datasets and assess how well our approach detects these anomalous occurrences.

This study provides an overview of abnormal activity detection, multiple face detection techniques, and recognition and classification algorithms.

II. METHODOLOGY

The methodology for abnormal event detection using CNN, AlexNet, VGG-16, and GoogleNet models involves several key steps aimed at effectively leveraging the capabilities of these deep learning architectures. Initially, the dataset for abnormal event detection needs to be prepared, comprising both normal and abnormal event instances across various contexts, such as surveillance videos or medical imaging. Data preprocessing techniques including normalization, resizing, and augmentation are applied to enhance model robustness and generalization.

Following data preparation, the selected deep learning models, namely CNN, AlexNet, VGG-16, and GoogleNet, are trained on the prepared dataset. During training, transfer learning may be employed, utilizing pre-trained weights on large-scale datasets like ImageNet to expedite convergence and enhance performance. Fine-tuning the model's parameters ensures adaptation to the specific characteristics of the abnormal event detection task.

After model training, evaluation is conducted to assess the performance of each architecture. Metrics such as accuracy, precision, recall, and F1-score are computed to gauge the models' ability to correctly identify abnormal events while minimizing false positives. Cross-validation techniques may be employed to ensure the reliability of the evaluation results.

In parallel, hyperparameter tuning is performed to optimize the performance of each model variant. Parameters such as learning rate, batch size, and dropout rates are systematically varied and validated using techniques like grid search or random search.

Once the models are trained and evaluated, they are deployed for real-time abnormal event detection. Deployment involves integrating the trained models into the target application environment, whether it's a surveillance system, medical diagnostic tool, or industrial monitoring system. Real-time inference is performed on incoming data streams, with detected abnormal events triggering appropriate alerts or actions.

Continuous monitoring and model maintenance are essential post-deployment activities to ensure sustained performance and adaptability to evolving abnormal event patterns, Regular retraining with updated data and periodic reassessment of model performance contribute to the long-term efficacy of the abnormal event detection.

The design phases of the method outlined in Figure 1 can be listed as follows:

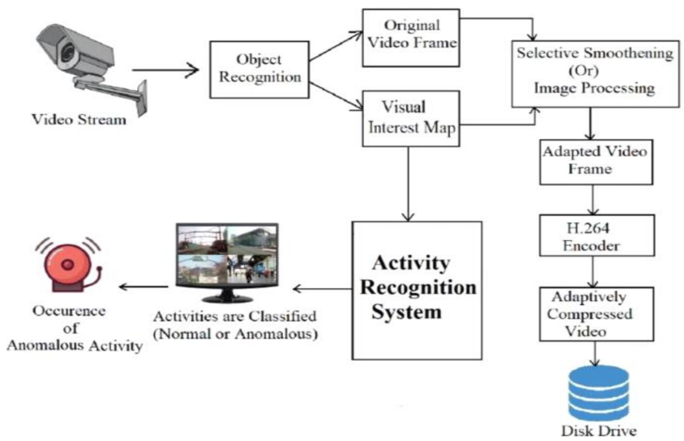

Fig. 1. Block diagram showing the outline of the abnormal event detection system.

MODEL TRAINING:

1. CNN (Convolutional Neural Network):

Convolutional neural network (CNN) is one type of Artificial Neural Network. A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and are used mainly for image processing, classification, segmentation and also for other auto correlated data.

- Design a CNN architecture suitable for abnormal event detection, considering factors like depth, filter sizes, and pooling layers.

- Train the CNN model on the training dataset.

- Use the validation set to tune hyper parameters and prevent over fitting.

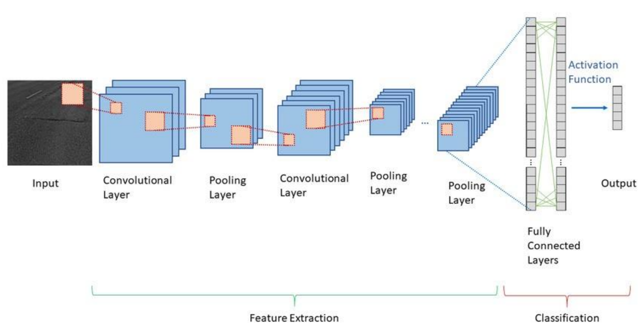

Fig.2. CNN Model

2. VGG16:

VGG stands for Visual Geometry Group it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. The "deep" refers to the number of layers with VGG-16 or VGG-19 consisting of 16 and 19 convolutional layers. The VGG architecture is the basis of ground-breaking object recognition models. Developed as a deep neural network, the VGGNet also surpasses baselines on many tasks and datasets beyond ImageNet. Moreover, it is now still one of the most popular image recognition architectures.

The VGG model, or VGGNet, that supports 16 layers is also referred to as VGG16, which is a convolutional neural network model proposed by A. Zisserman and K. Simonyan from the University of Oxford

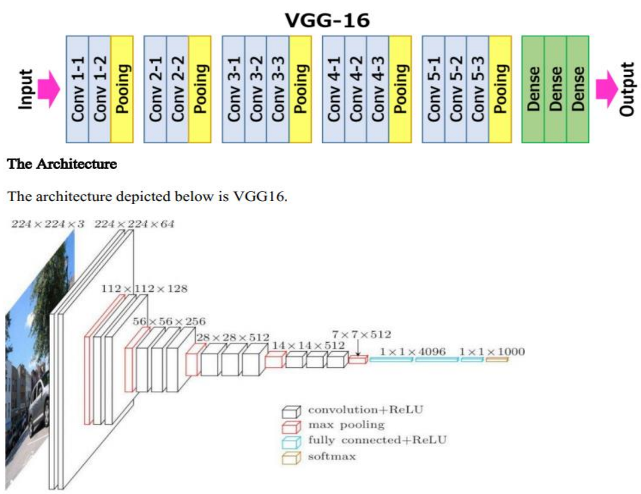

Fig.3. VGG16 Model

- Utilize the pre-trained VGG16 model with weights trained on ImageNet.

- Fine-tune the VGG16 model by adding additional layers suited for abnormal event detection.

- Train the modified VGG16 model on the training dataset and validate performance.

3. MobileNetV2:

- Utilize the lightweight MobileNetV2 architecture, which is efficient for deployment on resource-constrained devices.

- Fine-tune the MobileNetV2 model for abnormal event detection, potentially adjusting the number of layers or adding additional layers.

- Train the modified MobileNetV2 model on the training dataset and validate its performance.

4. AlexNet:

- Implement the AlexNet architecture, which was one of the pioneering deep CNN architectures.

- Adjust the AlexNet model architecture if necessary for abnormal event detection tasks.

- Train the modified AlexNet model on the training dataset and evaluate its performance.

MODEL EVALUATION:

1. Performance Metrics:
- Evaluate the models using appropriate performance metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve (AUC).
- Calculate these metrics on both the validation and test sets to ensure robustness.

2. Comparison:
- Compare the performance of CNN, VGG16, MobileNetV2, and AlexNet models in terms of their ability to detect abnormal events accurately.
- Identify which model performs best for the given dataset and task.

3. Cross-Validation:
- Perform cross-validation techniques such as k-fold cross-validation to validate the models' generalization ability and mitigate over fitting.

4. Deployment Considerations:
- Consider factors like model size, computational efficiency, and inference speed for deployment, especially if deploying on resource-constrained devices or real-time systems.

5. Further Analysis:
- Analyze misclassifications and errors made by the models to identify potential areas of improvement.
- Explore techniques such as ensemble learning or model stacking to further enhance detection performance.

Deep Learning Architectures:

To process the extracted features and detect abnormal events, we can employ CNNs, specifically, AlexNet and VGG16, due to their effectiveness in image classification tasks. We can fine-tune these pre-trained architectures on the extracted features to adapt them to the abnormal event detection task.

1. CNNs (e.g., AlexNet, VGG16):

- Utilize pre-trained CNN models to extract high-level features from the pre-processed video frames.

- Fine-tune these models on the extracted features to enhance their ability to discriminate between normal and abnormal events.

2. Hybrid Architectures:

- Explore hybrid architectures that combine CNNs with recurrent or attention mechanisms to capture both spatial and temporal dependencies in the video data.

Image Data Generator:

It is that rescales the image, applies shear in some range, zooms the image and does horizontal flipping with the image. This Image Data Generator includes all possible orientation of the image.

Training Process:Train_datagen.flow_from_directory is the function that is used to prepare data from the train_dataset directory Target_size specifies the target size of the image. Test_datagen.flow_from_directory is used to prepare test data for the model and all is similar as above. fit_generator is used to fit the data into the model made above, other factors used are steps_per_epochstells us about the number of times the model will execute for the training data.

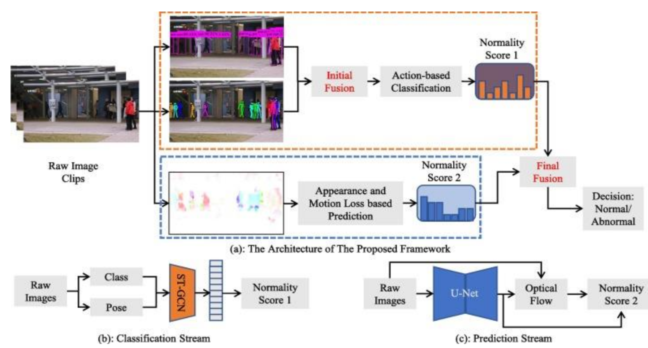

Fig.4. Training Dataset

III. MODELING AND ANALYSIS

MODEL DEVELOPMENT

Module Description: Abnormal Event Detection using Machine Learning with Adaptive Median Filter, Multiscale ROI Segmentation, CNN Classification, and VGG16 Analysis.

1. DATA COLLECTION

- Objective: Gather Comprehensive Dataset for Abnormal Event Detection.

1.1 Data Acquisition:

- Collect video data capturing various scenes where abnormal events might occur.
- Ensure diversity in environments, lighting conditions, and types of abnormal events.

1.2 Annotation and Labeling:

- Annotate collected video frames with labels indicating the presence or absence of abnormal events.
- Develop a standardized labeling protocol for consistent annotation across the dataset.

1.3 Data Augmentation:

- Apply augmentation techniques such as cropping, rotation, and brightness adjustment to diversify the dataset.
- Increase dataset size to enhance model generalization.

1.4 Secure Storage:

- Store collected and augmented datasets securely, adhering to data privacy regulations.
- Ensure accessibility while maintaining data confidentiality.

2. DATA PRE-PROCESSING

- Objective: Enhance Quality of Video Frames and Remove Noise

2.1 Noise Removal:

- Implement adaptive median filter to remove noise from video frames, preserving important features.
- Adjust filter parameters based on the characteristics of the input data.

2.2 Contrast Enhancement:

- Apply histogram equalization or contrast stretching to enhance visibility of abnormal events in video frames.

2.3 Normalization:

- Normalize pixel values to a standardized range to ensure consistency across frames.
- Resize frames to a uniform size for efficient processing.

2.4 Quality Assurance:

- Incorporate quality checks to evaluate the effectiveness of pre-processing techniques.
- Continuously refine pre-processing methods based on feedback.

3. DATA SEGMENTATION AND NORMALIZATION

- Objective: Extract Relevant Features for Abnormal Event Detection

3.1 Region of Interest Identification:

- Utilize multiscale region of interest (ROI) techniques to identify areas in video frames indicative of abnormal events.
- Capture variations in spatial and temporal characteristics across different scales.

3.2 Feature Extraction:

- Extract features from identified ROIs to represent spatial and temporal patterns associated with abnormal events.
- Include motion-based features, texture descriptors, and spatial relationships.

3.3 Normalization:

- Normalize extracted features to ensure consistency and improve model performance.

4. CLASSIFICATION (ABNORMAL EVENT DETECTION)

- Objective: Classify Video Frames to Detect Abnormal Events

4.1 CNN Model Training:

- Train a CNN classifier on pre-processed and segmented video frame features.
- Utilize labeled dataset for supervised learning to map features to abnormal event classes.

4.2 Event Classification:

- Classify video frames into discrete abnormal event classes using the trained CNN model.
- Determine probabilities of different abnormal events for each frame.

5. ANALYSIS (ABNORMAL EVENT DETECTION)

- Objective: Analyze Detected Events to Identify Abnormal Patterns

5.1 VGG16 Model Training:

- Train a VGG16 model on detected abnormal events to identify patterns indicative of specific abnormal scenarios.
- Utilize outputs from the CNN classifier for further analysis.

5.2 Performance Evaluation:

- Evaluate model performance using metrics such as accuracy, precision, recall, and F1 score.
- Analyze strengths and limitations of abnormal event detection model.

5.3 Feature Importance:

- Analyze importance of different features in abnormal event detection.
- Identify key indicators of abnormal events in spatial and temporal data.

6. PREDICTION AND ACTION

- Objective: Predict Abnormal Events and Take Action

6.1 Event Prediction:

- Predict abnormal events in real-time using the trained models.
- Categorize events into distinct categories based on severity or type.

6.2 Response Recommendation:

- Provide recommendations for appropriate responses or interventions based on predicted abnormal events.
- Trigger alerts, notifications, or automated actions to mitigate potential risks.

Import The Given Image From Dataset

We have to import our data set using keras preprocessing image data generator function also we create size, rescale, range, zoom range, horizontal flip. Then we import our image dataset from folder through the data generator function. Here we set train, test, and validation also we set target size, batch size and class-mode from this function we have to train using our own created network by adding layers of CNN.

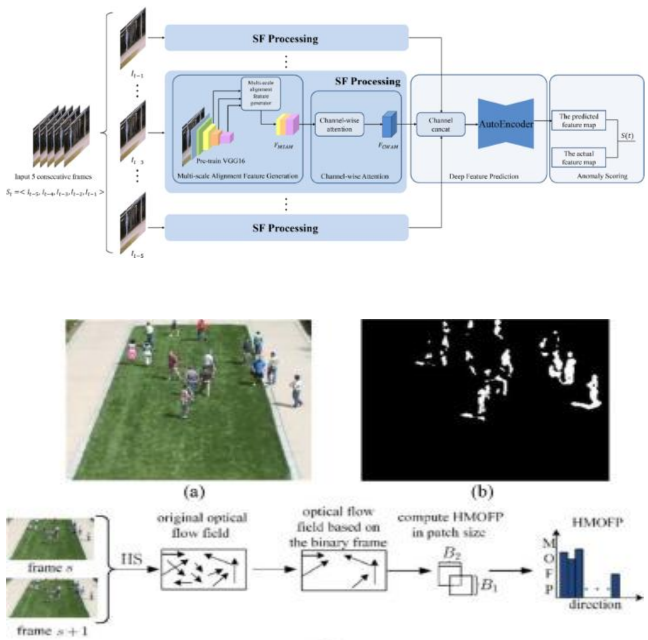

Fig. 5. Dataset Training

IV. RESULT AND DISCUSSION

This project presents a significant contribution to knowledge in the field of abnormal event detection through the exploration and comparison of convolutional neural network (CNN) architectures including VGG16, MobileNetV2, and AlexNet. By leveraging these architectures, the project establishes novel methodologies for identifying and classifying abnormal events within video streams with high accuracy and efficiency. The findings not only advance the understanding of deep learning-based anomaly detection but also offer practical insights for real-world applications such as surveillance, healthcare monitoring, and industrial safety.

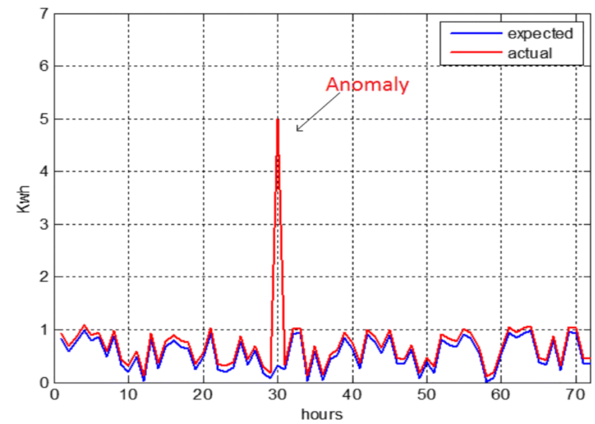

Fig.6. Trained Dataset Accuracy

V. CONCLUSION

In summary, convolutional neural networks (CNN) such as VGG16, MobileNetV2, and AlexNet show promise for future developments in computer vision applications when used for abnormal event detection. When it comes to identifying unusual actions or behaviors in video streams, these models are reliable and effective. Their capacity to decipher intricate visual data facilitates the improvement of public safety, surveillance systems, and several industrial processes. By incorporating these CNN designs into real-world scenarios and continuing research and development, there is great potential to improve anomaly detection skills and make a substantial contribution to security, safety, and efficiency in a variety of environments.

REFERENCES

[1]. J. Yu, Y. Lee, K. C. Yow, M. Jeon and W. Pedrycz, "Abnormal Event Detection and Localization via Adversarial Event Prediction," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 8, pp. 3572-3586, Aug. 2022, doi: 10.1109/TNNLS.2021.3053563.

[2]. .M. I. Georgescu, R. T. Ionescu, F. S. Khan, M. Popescu and M. Shah, "A Background-Agnostic Framework With Adversarial Training for Abnormal Event Detection in Video," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 9, pp. 4505-4523, 1 Sept. 2022, doi: 10.1109/TPAMI.2021.3074805.

[3]. J. Kwon and K. M. Lee, "A Unified Framework for Event Summarization and Rare Event Detection from Multiple Views," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1737-1750, 1 Sept. 2015, doi: 10.1109/TPAMI.2014.2385695.

[4]. G. Chen et al., "NeuroAED: Towards Efficient Abnormal Event Detection in Visual Surveillance With Neuromorphic Vision Sensor," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 923-936, 2021, doi: 10.1109/TIFS.2020.3023791.

[5]. Y. Yuan, Y. Feng and X. Lu, "Statistical Hypothesis Detector for Abnormal Event Detection in Crowded Scenes," in IEEE Transactions on Cybernetics, vol. 47, no. 11, pp. 3597-3608, Nov. 2017, doi: 10.1109/TCYB.2016.2572609.

[6]. W. Zhang, G. Wang, M. Huang, H. Wang and S. Wen, "Generative Adversarial Networks for Abnormal Event Detection in Videos Based on Self-Attention Mechanism," in IEEE Access, vol. 9, pp. 124847-124860, 2021, doi: 10.1109/ACCESS.2021.3110798.

[7]. S. CoÅŸar, G. Donatiello, V. Bogorny, C. Garate, L. O. Alvares and F. BrÃ©mond, "Toward Abnormal Trajectory and Event Detection in Video Surveillance," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 3, pp. 683-695, March 2017, doi: 10.1109/TCSVT.2016.2589859.

[8]. S. Yan, J. S. Smith, W. Lu and B. Zhang, "Abnormal Event Detection From Videos Using a Two-Stream Recurrent Variational Autoencoder," in IEEE Transactions on Cognitive and Developmental Systems, vol. 12, no. 1, pp. 30-42, March 2020, doi: 10.1109/TCDS.2018.2883368.

[9]. H. Song, C. Sun, X. Wu, M. Chen and Y. Jia, "Learning Normal Patterns via Adversarial Attention-Based Autoencoder for Abnormal Event Detection in Videos," in IEEE Transactions on Multimedia, vol. 22, no. 8, pp. 2138-2148, Aug. 2020, doi: 10.1109/TMM.2019.2950530.

[10]. S. Lee, H. G. Kim and Y. M. Ro, "BMAN: Bidirectional Multi-Scale Aggregation Networks for Abnormal Event Detection," in IEEE Transactions on Image Processing, vol. 29, pp. 2395-2408, 2020, doi: 10.1109/TIP.2019.2948286.

[11]. Y. Liu, J. Wang, J. Li, S. Niu, L. Wu and H. Song, "Zero-Bias Deep-Learning-Enabled Quickest Abnormal Event Detection in IoT," in IEEE Internet of Things Journal, vol. 9, no. 13, pp. 11385-11395, 1 July1, 2022, doi: 10.1109/JIOT.2021.3126819.

[12]. C. Sun, Y. Jia, H. Song and Y. Wu, "Adversarial 3D Convolutional Auto-Encoder for Abnormal Event Detection in Videos," in IEEE Transactions on Multimedia, vol. 23, pp. 3292-3305, 2021, doi: 10.1109/TMM.2020.3023303.

[13]. J. K. Dutta, B. Banerjee and C. K. Reddy, "RODS: Rarity based Outlier Detection in a Sparse Coding Framework," in IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 2, pp. 483-495, 1 Feb. 2016, doi: 10.1109/TKDE.2015.2475748.

[14]. O. Ye, J. Deng, Z. Yu, T. Liu and L. Dong, "Abnormal Event Detection via Feature Expectation Subgraph Calibrating Classification in Video Surveillance Scenes," in IEEE Access, vol. 8, pp. 97564-97575, 2020, doi: 10.1109/ACCESS.2020.2997357.

[15]. C. Huang, Z. Wu, J. Wen, Y. Xu, Q. Jiang and Y. Wang, "Abnormal Event Detection Using Deep Contrastive Learning for Intelligent Video Surveillance System," in IEEE Transactions on Industrial Informatics, vol. 18, no. 8, pp. 5171-5179, Aug. 2022, doi: 10.1109/TII.2021.3122801.

[16]. J. Yu, M. Kim, H. Oh and J. Yang, "Real-Time Abnormal Insider Event Detection on Enterprise Resource Planning Systems via Predictive Auto-Regression Model," in IEEE Access, vol. 9, pp. 62276-62284, 2021, doi: 10.1109/ACCESS.2021.3074149.

[17]. W. Chu, H. Xue, C. Yao and D. Cai, "Sparse Coding Guided Spatiotemporal Feature Learning for Abnormal Event Detection in Large Videos," in IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 246-255, Jan. 2019, doi: 10.1109/TMM.2018.2846411.

[18]. D. Ma, X. Hu, H. Zhang, Q. Sun and X. Xie, "A Hierarchical Event Detection Method Based on Spectral Theory of Multidimensional Matrix for Power System," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 4, pp. 2173-2186, April 2021, doi: 10.1109/TSMC.2019.2931316.

[19]. D. Tran, J. Yuan and D. Forsyth, "Video Event Detection: From Subvolume Localization to Spatiotemporal Path Search," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 2, pp. 404-416, Feb. 2014, doi: 10.1109/TPAMI.2013.137.

[20]. Y. Lee, D. K. Han and H. Ko, "Acoustic signal based abnormal event detection in indoor environment using multiclass adaboost," in IEEE Transactions on Consumer Electronics, vol. 59, no. 3, pp. 615-622, August 2013, doi: 10.1109/TCE.2013.6626247.

[21]. T. Wang and H. Snoussi, "Detection of Abnormal Visual Events via Global Optical Flow Orientation Histogram," in IEEE Transactions on Information Forensics and Security, vol. 9, no. 6, pp. 988-998, June 2014, doi: 10.1109/TIFS.2014.2315971.

[22]. B. Yu, Y. Liu and Q. Sun, "A Content-Adaptively Sparse Reconstruction Method for Abnormal Events Detection With Low-Rank Property," in IEEE Transactions on Systems,

Man, and Cybernetics: Systems, vol. 47, no. 4, pp. 704-716, April 2017, doi: 10.1109/TSMC.2016.2638048.