https://doi.org/10.48047/AFJBS.6.Si4.2024.3538-3553

African Journal of Biological Sciences

Anomaly Object Recognition In Surveillance Videos: A Review

Mrs Sangita Mahendra Rajput[1*], Prof .Dr Mangesh D Nikose[2]

[1*]Research Scholar, Department of Electrical and Electronics Engineering ,School of Engineering and Technology , Sandip University Nashik ,Maharashtra, 422213 India

[2]Professor & HOD, Department of Electrical and Electronics Engineering ,School of Engineering and Technology , Sandip University Nashik ,Maharashtra,422213 India

Article History
Volume 6, Issue Si4, 2024
Received: 15 April 2024.
Accepted: 15 July 2024
Published: 23 July 2024
doi:
10.48047/AFJBS.6.Si4.2024.3538-3553

Abstract–
This comprehensive literature survey investigates the evolving landscape of Anomaly Object Recognition in Surveillance Videos, focusing on advancements and methodologies proposed in recent research. In the realm of video surveillance, the escalating volume of visual data necessitates robust systems capable of autonomously detecting unusual events. The survey encompasses a diverse array of approaches, leveraging artificial intelligence, machine learning, and computer vision, with applications spanning public safety, security, and critical infrastructure protection.

The survey begins by outlining the critical importance of anomaly object recognition in automating the identification of deviations from expected patterns, addressing the limitations of traditional manual monitoring methods. Key components of anomaly object recognition systems, including learning normal patterns, real-time detection, alert generation, adaptability, and integration with surveillance infrastructure, are highlighted for a comprehensive understanding.

The literature survey subsequently categorizes and analyzes a plethora of research papers, each contributing unique perspectives to the field. The methodologies explored encompass high-dimensional classification modeling, optimization algorithms, generative networks, deep convolutional neural networks, and bidirectional consistency models. Applications range from public spaces and critical infrastructure to retail environments, traffic monitoring, and industrial security.

Key words: Anomaly Object Recognition, Surveillance Videos, Deep learning, machine learning

1 INTRODUCTION

In contemporary society, the ubiquity of surveillance systems has become an integral aspect of ensuring public safety, safeguarding critical infrastructure, and enhancing security across various domains. Surveillance videos serve as a critical source of information for monitoring public spaces, transportation systems, commercial establishments, and various other environments

As the volume of video data generated by surveillance cameras continues to escalate, the need for advanced techniques to decipher and interpret this wealth of information becomes imperative. Anomaly object recognition in surveillance videos stands at the forefront of cutting-edge research, aiming to automatically identify irregular and potentially threatening activities amidst the vast sea of visual data.

Anomaly object recognition is a critical aspect of video surveillance systems, aiming to identify unusual or unexpected events in real-world scenarios.

However, the sheer magnitude of data often overwhelms human operators, making it challenging to detect anomalous events or objects promptly. Anomaly object recognition, rooted in the realms of artificial intelligence (AI) and computer vision, seeks to alleviate this challenge by endowing surveillance systems with the capability to autonomously discern deviations from expected patterns or behaviors.

Anomaly object recognition, a subfield within video analytics, focuses on detecting unusual or abnormal objects or events in video streams

Video anomaly detection has garnered significant attention in recent years due to its crucial applications in surveillance, security, and safety. The literature reveals a growing interest in employing deep learning techniques for enhancing the accuracy and efficiency of anomaly detection systems. Themes identified in the literature include the utilization of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variants, such as LSTM and RCNN, for processing video data. Moreover, researchers have explored various architectures and methodologies, including single-stream and multi-stream networks, to capture spatial-temporal dependencies in videos.

Video anomaly detection is a crucial task in surveillance systems to identify unusual events or behaviors in video streams. With the advancement of deep learning techniques, various algorithms and methods have been developed to improve the accuracy and efficiency of anomaly detection. In this report, we will analyze the types of algorithms and methods used in recent research papers on video anomaly detection

1.1 Types of Algorithms

1..1.1 Deep Learning-Based Algorithms:

- Many recent studies employ deep learning architectures for video anomaly detection due to their ability to automatically learn hierarchical features from data.
- Common deep learning models include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and their variants.
- CNNs are typically used for spatial feature extraction, while RNNs and LSTMs are utilized for capturing temporal dependencies in video sequences.
- Examples of deep learning-based algorithms include Faster R-CNN, Generative Adversarial Networks (GANs), Autoencoders, and Transformer networks.

1.1.2 One-Class Learning Algorithms:

- One-Class Neural Networks (OCNNs) are designed to learn representations of normal data and detect deviations from this norm.
- These algorithms are trained on only normal instances and are capable of identifying anomalies as data points that deviate significantly from the learned normal behavior.

1.1.3 Hybrid and Ensemble Algorithms:

- Some studies combine multiple algorithms or models to leverage their complementary strengths.
- Hybrid approaches may combine deep learning with traditional machine learning algorithms or heuristic-based methods to enhance anomaly detection performance.
- Ensemble methods aggregate predictions from multiple models to improve robustness and generalization.

1.2 Methods Used:

1.2.1Spatio-Temporal Feature Extraction:

- Many algorithms focus on extracting spatio-temporal features from video data to capture both spatial patterns and temporal dynamics.
- Optical flow, motion vectors, histograms of oriented gradients (HOG), and deep feature representations are commonly used for this purpose.

1.2.2Memory Mechanisms:

- Memory-based methods, such as Memory Bank Modules, aim to memorize normal patterns and identify deviations from learned representations.
- Partitioned memory banks and error estimation techniques help in effectively capturing anomalies by comparing current observations with historical data.

1.2.3Self-Supervised and Multi-Task Learning:

- Self-supervised learning techniques train models to predict missing or corrupted data, encouraging the model to learn meaningful representations.
- Multi-task learning frameworks leverage auxiliary tasks to improve feature representations and anomaly detection performance.

1.2.4Attention Mechanisms:

- Attention mechanisms focus on relevant regions or frames within a video sequence, allowing the model to selectively attend to salient features.
- Attention-based CNN-LSTM structures and transformer networks have shown promising results in capturing anomalous events by attending to informative spatio-temporal regions.

Video anomaly detection research has seen significant advancements in recent years, primarily driven by deep learning techniques. Various algorithms and methods, including deep learning architectures, one-class learning algorithms, hybrid approaches, and ensemble methods, have been proposed to address the challenges of anomaly detection in videos. Incorporating spatio-temporal feature extraction, memory mechanisms, self-supervised learning, and attention mechanisms has contributed to improving the accuracy and robustness of anomaly detection systems. Further research in this area is expected to lead to more efficient and reliable video surveillance systems for real-world applications.

2 RELATED WORK

The following literature survey synthesizes key findings and methodologies from a diverse set of research papers addressing anomaly object recognition in surveillance videos.

2.1 Deep Learning-Based Approaches: Arunnehru [1] proposes a deep learning-based method for real-world object detection and improved anomaly detection. Leveraging advanced neural networks, the study contributes to enhancing the accuracy of anomaly detection in surveillance scenarios. Mansour et al. [2] present an intelligent video anomaly detection and classification system utilizing Faster R-CNN with deep reinforcement learning. The integration of deep learning and reinforcement learning techniques demonstrates effectiveness in accurately identifying anomalous events. Nawaratne et al. [3] focus on spatiotemporal anomaly detection using deep learning for real-time video surveillance. The study explores the capabilities of deep learning models in capturing complex temporal patterns for effective anomaly recognition .Zhou et al. [4] introduce AnomalyNet, a dedicated anomaly detection network for video surveillance. This research emphasizes the development of specialized architectures to enhance the accuracy of recognizing anomalous objects in surveillance footage. Pustokhina et al. [5] contribute to automated anomaly detection in pedestrian walkways. Their study employs automated deep learning-based methods to enhance safety through the recognition of anomalous behaviors of vulnerable road users.

2.2 Specialized Architectures and Techniques: Zhao et al. [6] propose a novel two-stream structure for video anomaly detection in smart city management. The research introduces specialized architectures tailored for anomaly detection in complex urban environments.
Wu et al. [7] present a deep one-class neural network designed for anomalous event detection in complex scenes. The study introduces a focused neural network architecture for robust anomaly recognition in challenging environments. Maqsood et al. [8] contribute to anomaly recognition from surveillance videos using a 3D convolutional neural network (CNN). The utilization of 3D CNNs enhances the model's ability to capture temporal dependencies, critical for anomaly detection. Gaus et al. [9] evaluate a dual convolutional neural network architecture for object-wise anomaly detection in cluttered X-ray security imagery. This research focuses on specialized architectures for object-level anomaly detection in security scenarios. Ingle and Kim [10] address real-time abnormal object detection for video surveillance in smart cities. Their work emphasizes the importance of real-time processing and detection capabilities for maintaining situational awareness in urban environments.

2.3 Integration of Multiple Modalities and Techniques: Boukabous and Azizi [11] contribute to image and video-based crime prediction using object detection and deep learning. The study explores the integration of object detection techniques with deep learning for crime prediction in surveillance scenarios. Xing and Li [12] propose visual anomaly detection via a partition memory bank module and error estimation. This research explores the integration of memory-based modules and error estimation techniques for improved anomaly detection. Lundström et al. [14] focus on improving deep learning-based anomaly detection on multivariate time series through separated anomaly scoring. Their approach integrates separated anomaly scoring techniques to enhance anomaly detection performance.

2.4 Survey and Review Papers: Hidayat [13] provides a systematic review of intelligent video analytics for suspicious object detection. This survey paper summarizes existing methodologies and approaches, offering insights into the landscape of video analytics for security applications. Nassif et al. [19] present a systematic review of machine learning techniques for anomaly detection. This survey paper provides a comprehensive overview of the state-of-the-art in machine learning methods applied to anomaly detection.

2.5Advanced Techniques and Innovations: Zhong et al. [15] characterize background-anomaly separability with a generative adversarial network for hyper spectral anomaly detection. The study introduces advanced techniques leveraging generative adversarial networks for improved hyper spectral anomaly detection. Lv et al. [16] focus on localizing anomalies from weakly-labeled videos. Their research contributes to the challenging task of precisely localizing anomalies within videos, especially when labeled data is limited. Wang et al. [17] propose unsupervised anomaly video detection via a Double-Flow ConvLSTM Variational Autoencoder. This work introduces a novel architecture for unsupervised anomaly detection, combining ConvLSTM and variational autoencoder techniques. Yan et al. [18] present unsupervised anomaly segmentation via multilevel image reconstruction and adaptive attention-level transition. The study focuses on unsupervised anomaly segmentation, emphasizing multilevel image reconstruction and adaptive attention mechanisms.

2.6 Recent Advances and Novel Frameworks: Luo et al. [21] introduce video anomaly detection with sparse coding-inspired deep neural networks. This research explores the integration of sparse coding techniques into deep neural networks for improved video anomaly detection. Castellani et al. [56] revisit self-supervised multi-task learning for video anomaly detection, introducing SSMTL++ as an advanced framework. Abdullah and Al-Ani [60] propose an adaptive algorithm based on principal component analysis-deep learning for anomalous events detection. Rendón-Segador et al. [61] present Crimenet, a neural structured learning approach using a vision transformer for violence detection.

2.7 Innovative Applications and Use Cases: Jezequel et al. [65] propose an efficient anomaly detection approach using self-supervised multi-cue tasks. The research focuses on incorporating multiple cues for effective anomaly recognition. AbuAlghanam et al. [66] introduce a fusion-based anomaly detection system using a modified isolation forest, specifically designed for the Internet of Things (IoT) applications.

Onyema et al. [63] present a remote monitoring system utilizing a slow-fast deep convolutional neural network model for identifying anti-social activities in surveillance applications. Simmers et al. [64] explore secure and intelligent video surveillance using unmanned aerial vehicles, providing a novel perspective on aerial surveillance applications.

2.8 Anomaly Detection

The section elaborates on the conventional methods of the research of video anomaly detection along with the benefits, and the challenges as follows,

In [71], Neda Tanghinezhad and Mehran Yazdi elaborated the video anomaly detection with the Multiscale-Multipath-U-Net (MSMP-Net), which works similarly to the encoder decoder structure of the U-Net. The model utilized certain unique modules such as the Multiscale-memorizer module and the Multi-resolution feature, which are the major contributions to obtaining the anomaly at the inconsistent size of the video with the minimal error map. The model, in addition, worked to improve the attention-driven scoring function of the model and enhanced the accuracy of the video anomaly detection. The hyperparameters should be adjusted to make the model adaptive, as the lack of adaptability resulted in poor performance when as comes to the real-time data.

Peng Xing, and Zechao Li [72] suggested the Partition Memory Bank (PMB) Mechanism in the research to learn the features that resulted in better performance in the detection of the video anomaly. Furthermore, the model included the Histogram Error Estimation Module (HEEM) that neglected the error that occurred in the PMB. The model ignored the most suitable error estimation models that were available currently along with the generalizability of the PMB model in the detection of the anomaly in the video.

Adam Lunderstrom*et al.*, [73] provided the autoencoder model along with the aggregated and the separated scoring method that detects the anomaly. The model provided better generalizability in the research as well and the results of the recall were highly determinable. The precision results obtained in the research were a bit lower than the other conventional methods and needed to be improved with the labeled data to obtain accurate results. The evaluation methods of the scoring and the respective algorithms could be separated and evaluated as the scoring mechanisms require the most significant attention.

In [74], the SSMTL++ model was introduced by Antonio Barbalau*et al.*, which included the YOLOv3 model to evaluate the anomaly object obtained in the video. Furthermore, the model utilized the 3D convolutional followed by the adversarial training, which showed improvement in the results that were correlated to the anomaly detection. The model undergo over fitting when worked with several proxy tasks that ignored the standard principle of anomaly detection in the videos.

Loic Jezequel*et al.* [75] in the self-supervised tasks suggested the proposed method that works based on the two branches of the network among which, one was the tint rotation detection task, and the other was the re-colorization task. The model needed to work with generative tasks that minimized the localized errors, and in addition, needed to train the model such as the OC-SVM that supports better detection of the anomaly.

In [76], Peng Wu *et al.* used the Deep One classifier Neural Network (deep OC-NN) to determine the anomalies in the video. The model Deep OC acted as the improved autoencoder to provide the high-level representations and in addition, acted as the lightweight and compatible model that worked faster to detect the anomalies. Moreover, the model included the multi-frame and optical flow that decided the spatiotemporal information of the objects. The model faced difficulties in detecting the objects and the movement of the anomaly that seemed to be distorted as the objects were too far from the camera. The model could work with generative models such as the Generative Adversarial Networks (GAN) that could destroy the blurriness and provide better temporal information about the anomaly found in the video.

In [77], Soheil Vosta and Kin-Choong Yow modeled the KianNet model that comprised the attention-based CNN-LSTM. The model obtained high results in the detection of the anomaly and acted as the most reliable model. The model showed an improvement in the results of the detection due to the inheritance of the dual attention mechanisms, but needed to be improved to make the model lightweight. The YOLOv3 model was utilized in the research to detect the movements of the objects in the video as well as the CNN-LSTM to classify the movements. The model could include the original frame along with the moving parts that were divided through the subtraction of the frames.

Jun-Fang Song *et al.* [78] presented the FusionNet-LSTM-G, which was based on the optical flow feature enhance Spatiotemporal feature Network. The model worked to extract the spatial and temporal features of the video. The model was the combination of the FusionNet-LSTM and the GAN that acted as the discriminator. The model should work on the de-shaking methods of the video and the updated target detection method to perform better anomaly detection in the video input.

The collection of papers presented here showcases the dynamic landscape of anomaly detection in video surveillance, spanning a rich spectrum of methodologies and innovations. From traditional approaches to state-of-the-art deep learning architectures, each paper contributes significantly to advancing the field's capabilities in detecting abnormal events and behaviors.

What stands out is the diverse range of techniques employed, including deep learning models like Faster R-CNN, convolutional and recurrent neural networks, as well as innovative frameworks integrating reinforcement learning, generative adversarial networks, and attention mechanisms. These approaches demonstrate a keen understanding of the complexities involved in real-world surveillance scenarios and strive for robustness, efficiency, and scalability.

Moreover, the papers reflect a growing emphasis on addressing specific challenges such as spatiotemporal modeling, background-foreground separation, multiscale feature fusion, and domain adaptation, underscoring the commitment of researchers to push the boundaries of anomaly detection performance across various application domains.

In addition to technical contributions, several papers delve into practical considerations such as dataset availability, evaluation metrics, and real-world deployment, fostering a holistic understanding of the challenges and opportunities in deploying anomaly detection systems.

3 RESEARCH GAPS

In the subsequent sections of this literature review, we delve deeper into these challenges, exploring existing research efforts, identifying gaps in the literature, and proposing avenues for future research to address these challenges comprehensively.

Anomaly detection and object detection are pivotal tasks in computer vision with far-reaching applications in surveillance, security, and autonomous systems. However, these tasks are fraught with challenges that impede their effectiveness and reliability. We delve into the intricate landscape of anomaly detection and object detection, highlighting the multifaceted challenges that researchers and practitioners encounter in these domains.

Anomaly detection, the process of identifying patterns in data that deviate from normal behavior, confronts several formidable hurdles. One of the foremost challenges lies in appropriate feature extraction, where the extraction of discriminative features from complex and high-dimensional data remains a daunting task. Moreover, defining normal behaviors in diverse contexts poses a significant challenge, as the concept of normalcy can vary widely across different domains and environments. Compounding this issue is the intricate task of accurately classifying data as normal or abnormal, which often entails addressing the variations and nuances inherent in abnormal behavior. Furthermore, the sporadic occurrence of abnormal events, coupled with environmental variations and camera movements, exacerbates the difficulty of anomaly detection, necessitating robust and adaptive detection mechanisms.

Similarly, object detection, the process of locating and classifying objects within an image or video, presents its own set of challenges. Object detection is often considered more arduous than image classification due to several factors. Dual priorities, where the algorithm must simultaneously focus on both localization and classification tasks, pose a significant computational burden. Moreover, achieving real-time performance while maintaining accuracy remains a persistent challenge, especially in scenarios with multiple scales, limited data, and class imbalance. Additionally, occlusion of objects, where only a fraction of an object may be visible, adds complexity to the detection process, requiring sophisticated algorithms capable of reasoning about occluded objects. Furthermore, objects blending into the background further compound the challenge, demanding robust detection algorithms capable of discerning subtle visual cues amidst cluttered backgrounds.

To summarize the challenges and the gap in research in bullet points we can say

- Challenges in anomaly detection include appropriate feature extraction, defining normal behaviours, classification of normal and abnormal data, addressing the variations in abnormal behaviour, sparse occurrence of abnormal events, environmental variations, and camera movements.
- Object detection is normally considered to be much harder than image classification ,particularly because of challenges like dual priorities, speed, multiple scales, limited data, and class imbalance.

- The objects of interest can be occluded, which means sometimes only a small portion of an object may be visible making it challenging to detect.
- Another common challenge is that the objects of interest may blend into the background, making them hard to identify.

In light of these challenges, researchers and practitioners in the fields of anomaly detection and object detection are tasked with developing innovative solutions that transcend traditional methodologies. Addressing these research gaps requires interdisciplinary collaboration, leveraging advancements in deep learning, computer vision, and pattern recognition. By surmounting these challenges, we can unlock the full potential of anomaly detection and object detection technologies, ushering in a new era of intelligent surveillance, security, and automation.

4.CHALLENGES AND FUTURE DIRECTIONS:

Future research endeavors should focus on addressing the identified research gaps by investigating novel deep learning architectures, incorporating multimodal data fusion techniques, and exploring alternative training strategies such as transfer learning and domain adaptation. Additionally, researchers should emphasize the development of benchmark datasets and evaluation protocols to facilitate fair comparisons between different anomaly detection methods and promote reproducibility in the field.

Several papers address the challenges associated with anomaly detection in real-world scenarios, including cluttered environments, limited labeled data, and dynamic backgrounds.

4.1 Challenges

Here are some categories of challenges commonly addressed in the literature:

4.1.1 Model Complexity and Scalability: Many papers deal with the challenge of designing models that are both effective in detecting anomalies and scalable to handle large-scale surveillance systems. This involves optimizing computational resources and memory usage to ensure real-time processing of video streams.

4.1.2 Feature Representation and Learning: Developing robust feature representations that capture both spatial and temporal information from video data is a key challenge. Papers address this by exploring various deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variants, to extract informative features.

4.1.3 Anomaly Definition and Labeling: Defining anomalies and obtaining labeled data for training models is often challenging. Papers explore different types of anomalies, ranging from rare events to subtle abnormalities, and propose methods to annotate data or leverage weakly supervised learning techniques to overcome the scarcity of labeled examples.

4.1.4 Model Generalization and Adaptability: Ensuring that anomaly detection models generalize well across different environments and are adaptable to changing conditions is crucial. This involves domain adaptation techniques, transfer learning approaches, and methods for fine-tuning models on specific target domains.

4.1.5 Real-world Deployment and Integration: Deploying anomaly detection systems in real-world surveillance scenarios poses challenges related to system integration, interoperability with existing infrastructure, and addressing practical constraints such as power consumption and hardware limitations.

4.1.6 Interpretability and Explainability: Interpreting the decisions made by anomaly detection models is essential for building trust in the system and enabling human operators to take appropriate actions. Papers investigate methods for explaining model predictions and providing meaningful insights into detected anomalies.

4.1.7 Data Imbalance and Noise: Imbalanced datasets with a scarcity of anomalous examples and noise in surveillance videos can hinder the performance of anomaly detection models. Techniques such as data augmentation, anomaly oversampling, and robust loss functions are explored to mitigate these challenges.

4.1.8 Temporal Context Modeling: Modeling temporal dependencies and long-range dependencies in video sequences is critical for accurately detecting anomalies. Papers propose various architectures, such as recurrent neural networks (RNNs), attention mechanisms, and spatio-temporal convolutional networks, to effectively capture temporal context information.

4.1.9 Evaluation Metrics and Benchmarking: Assessing the performance of anomaly detection algorithms requires appropriate evaluation metrics and benchmark datasets. Papers investigate different evaluation protocols, propose novel metrics tailored to specific application scenarios, and contribute annotated datasets for benchmarking purposes.

4.1.10 Privacy and Ethical Considerations: Addressing privacy concerns and ensuring the ethical use of surveillance data are important considerations in deploying anomaly detection systems. Papers discuss privacy-preserving techniques, anonymization methods, and ethical guidelines for collecting and analyzing surveillance data.
These categories highlight the diverse range of challenges tackled by researchers in the field of video anomaly detection, reflecting the interdisciplinary nature of this area encompassing computer vision, machine learning, and security.

Future research directions include the exploration of explainable AI techniques for anomaly detection, improved generalization across diverse scenarios, and the development of scalable frameworks for extremely large-scale video datasets.

4.2 Future Scope

To provide an extensive future scope for the surveyed papers, the outline of potential directions and advancements in the field of video anomaly detection are generalized as below:

4.2.1 Integration of Multiple Modalities: Future research could explore the integration of multiple modalities such as audio, depth, and thermal data alongside visual information to enhance anomaly detection accuracy and robustness. Fusion techniques, including attention mechanisms and multimodal learning architectures, could be employed for effective integration.

4.2.2 Continual Learning and Adaptation: Developing anomaly detection systems capable of continual learning and adaptation to changing environments is essential. Techniques such as online learning, transfer learning, and meta-learning could enable models to adapt to new anomalies and scenarios without requiring extensive retraining.

4.2.3 Spatiotemporal Context Modeling: Enhancing models' ability to capture complex spatiotemporal patterns and context is crucial. Future research could focus on developing advanced architectures, such as graph-based models or attention mechanisms over spatiotemporal graphs, to better understand the context of anomalies in video sequences.

4.2.4 Explainable AI: Providing interpretability and explainability in anomaly detection systems is increasingly important, especially in critical applications such as surveillance and security. Future research could explore techniques for generating human-understandable explanations for detected anomalies, enabling end-users to trust and verify system decisions.

4.2.5 _Robustness to Adversarial Attacks: Investigating the robustness of anomaly detection models against adversarial attacks is essential for real-world deployment. Future work could focus on developing adversarially robust architectures and training strategies to mitigate the impact of adversarial perturbations on model performance.

4.2.6 Real-time and Edge Deployment: Optimizing anomaly detection algorithms for real-time and edge deployment is crucial for applications where low latency and resource efficiency are paramount. Future research could explore lightweight architectures, quantization techniques, and hardware-accelerated solutions to enable efficient inference on edge devices.

4.2.7 Privacy-Preserving Techniques: Addressing privacy concerns while performing anomaly detection in surveillance videos is important. Future work could investigate privacy-preserving techniques, such as federated learning, differential privacy, and encrypted inference, to protect sensitive information while still enabling effective anomaly detection.

4.2.8 Long-Term Anomaly Forecasting: Moving beyond traditional anomaly detection, future research could explore the prediction and forecasting of long-term anomalies. By analyzing historical data and trends, models could anticipate future anomalies, enabling proactive decision-making and risk mitigation strategies.

4.2.9 Benchmark Datasets and Evaluation Metrics: Continued efforts in developing comprehensive benchmark datasets and standardized evaluation metrics are essential for advancing the field of video anomaly detection. Future research could focus on curating diverse datasets with varying complexities and defining evaluation protocols that capture real-world performance accurately.

4.2.10 Ethical and Societal Implications: As anomaly detection systems become more pervasive, addressing ethical and societal implications is crucial. Future research should consider the ethical use of surveillance technologies, potential biases in algorithmic decision-making, and the impact of automation on privacy and civil liberties.

By focusing on these future directions, researchers can contribute to the development of more reliable, interpretable, and ethical video anomaly detection systems with broader applicability across various domains and industries.

5. SUMMARY AND CONCLUSION

The collection of papers presented here constitutes a comprehensive exploration of anomaly detection and object detection in video surveillance, leveraging advanced techniques such as deep learning, reinforcement learning, and innovative network architectures. These papers address various challenges in anomaly detection and object detection, including occlusion, background clutter, class imbalance, and real-time performance requirements. Below is an extensive summary and conclusion drawn from the key findings and contributions of these papers

Several papers in this collection focus on anomaly detection in surveillance videos, aiming to identify events or behaviors that deviate from normal patterns. Techniques employed include deep learning models, such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Generative Adversarial Networks (GANs), among others. Noteworthy contributions include:

- Arunnehru (2023) proposes a deep learning-based approach for real-world object detection and improved anomaly detection. The study demonstrates the effectiveness of the proposed method in enhancing surveillance video analysis by accurately detecting anomalies.
- Mansour et al. (2021) introduce an intelligent video anomaly detection and classification framework using Faster R-CNN with deep reinforcement learning. Their model achieves superior performance in detecting and classifying anomalies in surveillance videos.
- Zhou et al. (2019) present AnomalyNet, a dedicated anomaly detection network for video surveillance. The network effectively captures spatiotemporal patterns in surveillance videos, enabling robust anomaly detection performance.

Object detection is another critical aspect of video surveillance, facilitating the identification and localization of objects of interest within a scene. The papers in this collection propose various methodologies to address challenges such as occlusion, scale variation, and background clutter. Key contributions include:

- Gaus et al. (2019) evaluate a dual convolutional neural network architecture for object-wise anomaly detection in cluttered X-ray security imagery. Their approach demonstrates promising results in detecting anomalous objects amidst cluttered backgrounds.
- Ingle and Kim (2022) propose a real-time abnormal object detection system for video surveillance in smart cities. Their system leverages deep learning techniques to accurately detect and localize abnormal objects, contributing to enhanced safety and security in urban environments.

The literature survey highlights the evolution of anomaly object recognition in surveillance videos, encompassing a wide array of techniques, architectures, and applications. From deep learning-based approaches to specialized network architectures, the surveyed papers collectively contribute to advancing the state-of-the-art in anomaly detection, addressing challenges and opening new avenues for research in this critical domain of video analytics for security and safety. While deep learning-based approaches have shown promise in video anomaly detection, further research is needed to overcome existing challenges and advance the state-of-the-art in this critical domain of surveillance and security.

In conclusion, the papers presented in this collection underscore the importance of advanced techniques in addressing the challenges of anomaly detection and object detection in video surveillance. The adoption of deep learning models, reinforcement learning, and innovative network architectures has significantly advanced the state-of-the-art in these domains, enabling more accurate, efficient, and reliable surveillance systems. However, several research directions warrant further exploration, including the development of robust algorithms for handling occlusion, scale variation, and background clutter, as well as the integration of multimodal data sources for enhanced anomaly detection capabilities. Overall, the findings and methodologies presented in these papers pave the way for future advancements in video surveillance technology, with profound implications for safety, security, and public welfare.

REFERENCES

[1] Arunnehru, J. "Deep learning-based real-world object detection and improved anomaly detection for surveillance videos." Materials Today: Proceedings 80 (2023): 2911-2916.

[2] Mansour, Romany F., José Escorcia-Gutierrez, Margarita Gamarra, Jair A. Villanueva, and Nallig Leal. "Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model." Image and Vision Computing 112 (2021): 104229.

[3] Nawaratne, Rashmika, Damminda Alahakoon, Daswin De Silva, and Xinghuo Yu. "Spatiotemporal anomaly detection using deep learning for real-time video surveillance." IEEE Transactions on Industrial Informatics 16, no. 1 (2019): 393-402.

[4] Zhou, Joey Tianyi, Jiawei Du, Hongyuan Zhu, Xi Peng, Yong Liu, and Rick Siow Mong Goh. "Anomalynet: An anomaly detection network for video surveillance." IEEE Transactions on Information Forensics and Security 14, no. 10 (2019): 2537-2550.

[5] Pustokhina, Irina V., Denis A. Pustokhin, Thavavel Vaiyapuri, Deepak Gupta, Sachin Kumar, and K. Shankar. "An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety." Safety Science 142 (2021): 105356.

[6] Zhao, Yuxuan, Ka Lok Man, Jeremy Smith, and Sheng-Uei Guan. "A novel two-stream structure for video anomaly detection in smart city management." Springer The Journal of Supercomputing 78, no. 3 (2022): 3940-3954.

[7] Wu, Peng, Jing Liu, and Fang Shen. "A deep one-class neural network for anomalous event detection in complex scenes." IEEE Transactions on neural networks and learning systems 31, no. 7 (2019): 2609-2622.

[8] [8] Maqsood, Ramna, 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 80, no. 12 (2021): 18693-18716.

[9] [9] Gaus, Yona Falinie A., Neelanjan Bhowmik, Samet Akçay, Paolo M. Guillén-Garcia, Jack W. Barker, and Toby P. Breckon. "Evaluation of a dual convolutional neural network architecture for object-wise anomaly detection in cluttered X-ray security imagery." In 2019 international joint conference on neural networks (IJCNN), pp. 1-8. IEEE, 2019.

[10] Ingle, Palash Yuvraj, and Young-Gab Kim. "Real-time abnormal object detection for video surveillance in smart cities." Sensors 22, no. 10 (2022): 3862.

[11] Boukabous, Mohammed, and Mostafa Azizi. "Image and video-based crime prediction using object detection and deep learning." Bulletin of Electrical Engineering and Informatics 12, no. 3 (2023): 1630-1638.

[12] Xing, Peng, and Zechao Li. "Visual anomaly detection via partition memory bank module and error estimation." IEEE Transactions on Circuits and Systems for Video Technology (2023).

[13] Hidayat, Fadhil. "Intelligent video analytic for suspicious object detection: a systematic review." In 2020 International Conference on ICT for Smart Society (ICISS), pp. 1-8. IEEE, 2020.

[14] Lundström, Adam, Mattias O'nils, Faisal Z. Qureshi, and Axel Jantsch. "Improving deep learning based anomaly detection on multivariate time series through separated anomaly scoring." IEEE Access 10 (2022): 108194-108204.

[15] Zhong, Jiaping, Weiying Xie, Yunsong Li, Jie Lei, and Qian Du. "Characterization of background-anomaly separability with generative adversarial network for hyperspectral anomaly detection." IEEE Transactions on Geoscience and Remote Sensing 59, no. 7 (2020): 6017-6028.

[16] H. Lv, C. Zhou, Z. Cui, C. Xu, Y. Li and J. Yang, "Localizing Anomalies From Weakly-Labeled Videos," in *IEEE Transactions on Image Processing*, vol. 30, pp. 4505-4515, 2021.

[17] L. Wang, H. Tan, F. Zhou, W. Zuo and P. Sun, "Unsupervised Anomaly Video Detection via a Double-Flow ConvLSTM Variational Autoencoder," in *IEEE Access*, vol. 10, pp. 44278-44289, 2022.

[18] Yan, Yi, Deming Wang, Guangliang Zhou, and Qijun Chen. "Unsupervised anomaly segmentation via multilevel image reconstruction and adaptive attention-level transition." IEEE Transactions on Instrumentation and Measurement 70 (2021): 1-12.

[19] Nassif, Ali Bou, Manar Abu Talib, Qassim Nasir, and Fatima Mohamad Dakalbab. "Machine learning for anomaly detection: A systematic review." Ieee Access 9 (2021): 78658-78700.

[20] Wang, Wenzheng, Wei Song, Zhen Li, Boya Zhao, and Baojun Zhao. "A novel filter-based anomaly detection framework for hyperspectral imagery." IEEE Access 9 (2021): 124033-124043.

[21] Luo, Weixin, Wen Liu, Dongze Lian, Jinhui Tang, Lixin Duan, Xi Peng, and Shenghua Gao. "Video anomaly detection with sparse coding inspired deep neural networks." IEEE transactions on pattern analysis and machine intelligence 43, no. 3 (2019): 1070-1084.

[22] Castellani, Andrea, Sebastian Schmitt, and Stefano Squartini. "Real-world anomaly detection by using digital twin systems and weakly supervised learning." IEEE Transactions on Industrial Informatics 17, no. 7 (2021): 4733-4742.

[23] Yahyaoui, Aymen, Takoua Abdellatif, Sami Yangui, and Rabah Attia. "READ-IoT: reliable event and anomaly detection framework for the Internet of Things." IEEE Access 9 (2021): 24168-24186.

[24] Jin, Pu, Lichao Mou, Gui-Song Xia, and Xiao Xiang Zhu. "Anomaly detection in aerial videos with transformers." IEEE Transactions on Geoscience and Remote Sensing 60 (2022): 1-13.

[25] Wang, Xuanzhao, Zhengping Che, Bo Jiang, Ning Xiao, Ke Yang, Jian Tang, Jieping Ye, Jingyu Wang, and Qi Qi. "Robust unsupervised video anomaly detection by multipath frame prediction." IEEE transactions on neural networks and learning systems 33, no. 6 (2021): 2301-2312.

[26] Song, Jun-Fang, Hai-Li Zhao, Duo-Yang Wen, and Xiao-Yu Xu. "Video Anomaly Detection Based on Optical Flow Feature Enhanced Spatio-Temporal Feature Network FusionNet-LSTM-G." IEEE Access 10 (2022): 130314-130325.

[27] Yuan, Hongchun, Zhenyu Cai, Hui Zhou, Yue Wang, and Xiangzhi Chen. "Transanomaly: Video anomaly detection using video vision transformer." IEEE Access 9 (2021): 123977-123986.

[28] Chen, Yurong, Hui Zhang, Yaonan Wang, Yimin Yang, Xianen Zhou, and QM Jonathan Wu. "MAMA Net: Multi-scale attention memory autoencoder network for anomaly detection." IEEE Transactions on Medical Imaging 40, no. 3 (2020): 1032-1041.

[29] Tariq, Sameema, Haroon Farooq, Abdul Jaleel, and Syed Muhammad Wasif. "Anomaly detection with particle filtering for online video surveillance." IEEE Access 9 (2021): 19457-19468.

[30] Mehmood, Abid. "Efficient anomaly detection in crowd videos using pre-trained 2D convolutional neural networks." IEEE Access 9 (2021): 138283-138295.

[31] Dataset - http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm - (Accessed on August 2023)

[32] Algamal, Zakariya Yahya, Maimoonah Khalid Qasim, Muhammad Hisyam Lee, and Haithem Taha Mohammad Ali. "High-dimensional QSAR/QSPR classification modeling based on improving pigeon optimization algorithm." Chemometrics and Intelligent Laboratory Systems 206 (2020): 104170.

[33] Wang, Feng, Heng Zhang, and Aimin Zhou. "A particle swarm optimization algorithm for mixed-variable optimization problems." Swarm and Evolutionary Computation 60 (2021): 100808.

[34] Yang, Zhiwei, Jing Liu, and Peng Wu. "Bidirectional retrospective generation adversarial network for anomaly detection in videos." IEEE Access 9 (2021): 107842-107857.

[35] Saypadith, Savath, and Takao Onoye. "An approach to detect anomaly in video using deep generative network." IEEE Access 9 (2021): 150903-150910.

[36] Fang, Zhiwen, Jiafei Liang, Joey Tianyi Zhou, Yang Xiao, and Feng Yang. "Anomaly detection with bidirectional consistency in videos." IEEE transactions on neural networks and learning systems 33, no. 3 (2020): 1079-1092.

[37] Saypadith, Savath, and Takao Onoye. "An approach to detect anomaly in video using deep generative network." IEEE Access 9 (2021): 150903-150910.

[38] Tariq, Sameema, Haroon Farooq, Abdul Jaleel, and Syed Muhammad Wasif. "Anomaly detection with particle filtering for online video surveillance." IEEE Access 9 (2021): 19457-19468.

[39] Fang, Zhiwen, Jiafei Liang, Joey Tianyi Zhou, Yang Xiao, and Feng Yang. "Anomaly detection with bidirectional consistency in videos." IEEE transactions on neural networks and learning systems 33, no. 3 (2020): 1079-1092.

[40] Mehmood, Abid. "Efficient anomaly detection in crowd videos using pre-trained 2D convolutional neural networks." IEEE Access 9 (2021): 138283-138295.

[41] Pillai, Gargi V., and Debashis Sen. "Anomaly detection in nonstationary videos using time-recursive differencing network-based prediction." IEEE Geoscience and Remote Sensing Letters 19 (2021): 1-5.

[42] Fang, Le, Yu Wu, Chuan Wu, and Yizhou Yu. "A nonintrusive elderly home monitoring system." IEEE Internet of Things journal 8, no. 4 (2020): 2603-2614.

[43] Hirasawa, Kaito, Keisuke Maeda, Takahiro Ogawa, and Miki Haseyama. "Detection of Important Scenes in Baseball Videos via Bidirectional Time Lag Aware Deep Multiset Canonical Correlation Analysis." IEEE Access 9 (2021): 84971-84981.

[44] Wang, Guodong, Yunhong Wang, Jie Qin, Dongming Zhang, Xiuguo Bao, and Di Huang. "Video anomaly detection by solving decoupled spatio-temporal jigsaw puzzles." In European Conference on Computer Vision, pp. 494-511. Cham: Springer Nature Switzerland, 2022.

[45] Hao, Yi, Jie Li, Nannan Wang, Xiaoyu Wang, and Xinbo Gao. "Spatiotemporal consistency-enhanced network for video anomaly detection." Elsevier Pattern Recognition 121 (2022): 108232.

[46] Li, Guoqiu, Guanxiong Cai, Xingyu Zeng, and Rui Zhao. "Scale-aware spatio-temporal relation learning for video anomaly detection." In European Conference on Computer Vision, pp. 333-350. Cham: Springer Nature Switzerland, 2022.

[47] Ganokratanaa, Thittaporn, Supavadee Aramvith, and Nicu Sebe. "Video anomaly detection using deep residual-spatiotemporal translation network." Elsevier Pattern Recognition Letters 155 (2022): 143-150.

[48] Cho, MyeongAh, Taeoh Kim, Woo Jin Kim, Suhwan Cho, and Sangyoun Lee. "Unsupervised video anomaly detection via normalizing flows with implicit latent features." Elsevier Pattern Recognition 129 (2022): 108703.

[49] Kamoona, Ammar Mansoor, Amirali Khodadadian Gostar, Alireza Bab-Hadiashar, and Reza Hoseinnezhad. "Multiple instance-based video anomaly detection using deep temporal encoding-decoding." Elsevier Expert Systems with Applications 214 (2023): 119079.

[50] Chen, Fanglin, Weihang Wang, Huiyuan Yang, Wenjie Pei, and Guangming Lu. "Multiscale feature fusion for surveillance video diagnosis." Elsevier Knowledge-Based Systems 240 (2022): 108103.

[51] Abdullah, Faisal, and Ahmad Jalal. "Semantic segmentation based crowd tracking and anomaly detection via neuro-fuzzy classifier in smart surveillance system." Springer Arabian Journal for Science and Engineering 48, no. 2 (2023): 2173-2190.

[52] Qi, Zhenting, Ruike Zhu, Zheyu Fu, Wenhao Chai, and Volodymyr Kindratenko. "Weakly Supervised Two-Stage Training Scheme for Deep Video Fight Detection Model." In 2022 IEEE

34th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 677-685. IEEE, 2022.

[53] Sabih, Mohammad, and Dinesh Kumar Vishwakarma. "Crowd anomaly detection with LSTMs using optical features and domain knowledge for improved inferring." Springer The Visual Computer 38, no. 5 (2022): 1719-1730.

[54] Aslam, Nazia, and Maheshkumar H. Kolekar. "Unsupervised anomalous event detection in videos using spatio-temporal inter-fused autoencoder." Springer Multimedia Tools and Applications 81, no. 29 (2022): 42457-42482.

[55] Barbalau, Antonio, Radu Tudor Ionescu, Mariana-Iuliana Georgescu, Jacob Dueholm, Bharathkumar Ramachandra, Kamal Nasrollahi, Fahad Shahbaz Khan, Thomas B. Moeslund, and Mubarak Shah. "SSMTL++: Revisiting self-supervised multi-task learning for video anomaly detection." Elsevier Computer Vision and Image Understanding 229 (2023): 103656.

[56] Wang, Zhongyue, and Ying Chen. "Anomaly detection with dual-stream memory network." Elsevier Journal of Visual Communication and Image Representation 90 (2023): 103739.

[57] V. D. Huszár, V. K. Adhikarla, I. Négyesi and C. Krasznay, "Toward Fast and Accurate Violence Detection for Automated Video Surveillance Applications," in *IEEE Access*, vol. 11, pp. 18772-18793, 2023, doi: 10.1109/ACCESS.2023.

[58] [59] S. Pouyan, M. Charmi, A. Azarpeyvand and H. Hassanpoor, "Propounding First Artificial Intelligence Approach for Predicting Robbery Behavior Potential in an Indoor Security Camera," in *IEEE Access*, vol. 11, pp. 60471-60489, 2023, doi: 10.1109/ACCESS.2023.

[59] [60] Abbas, Zainab K., and Ayad A. Al-Ani. "An adaptive algorithm based on principal component analysis-deep learning for anomalous events detection." Indonesian Journal of Electrical Engineering and Computer Science 29, no. 1 (2023): 421-430.

[60] [61] Rendón-Segador, Fernando J., Juan A. Álvarez-García, Jose L. Salazar-González, and Tatiana Tommasi. "Crimenet: Neural structured learning using vision transformer for violence detection." Elsevier Neural networks 161 (2023): 318-329.

[61] Pazho, Armin Danesh, Christopher Neff, Ghazal Alinezhad Noghre, Babak Rahimi Ardabili, Shanle Yao, Mohammadreza Baharani, and Hamed Tabkhi. "Ancilia: Scalable intelligent video surveillance for the artificial intelligence of things." IEEE Internet of Things Journal (2023).

[62] Onyema, Edeh Michael, Sundaravadivazhagn Balasubaramanian, Celestine Iwendi, BVV Siva Prasad, and Chinecherem Deborah Edeh. "Remote monitoring system using slow-fast deep convolution neural network model for identifying anti-social activities in surveillance applications." Elsevier Measurement: Sensors 27 (2023): 100718.

[63] Simmers, Ethan, Ahmad Salman, Evan Day, and Alma Oracevic. "Secure and Intelligent Video Surveillance using Unmanned Aerial Vehicles." In 2023 IEEE International Conference on Smart Mobility (SM), pp. 98-103. IEEE, 2023.

[64] Jezequel, Loic, Ngoc-Son Vu, Jean Beaudet, and Aymeric Histace. "Efficient anomaly detection using self-supervised multi-cue tasks." IEEE Transactions on Image Processing 32 (2023): 807-821.

[65] AbuAlghanam, Orieb, Hadeel Alazzam, Esra'A. Alhenawi, Mohammad Qatawneh, and Omar Adwan. "Fusion-based anomaly detection system using modified isolation forest for internet of things." Spinger Journal of Ambient Intelligence and Humanized Computing 14, no. 1 (2023): 131-145.

[66] Noh, DongKi, Changki Sung, Teayoung Uhm, WooJu Lee, Hyungtae Lim, Jaeseok Choi, Kyuewang Lee et al. "X-MAS: Extremely Large-Scale Multi-Modal Sensor Dataset for Outdoor

Surveillance in Real Environments." IEEE Robotics and Automation Letters 8, no. 2 (2023): 1093-1100.

[67] Mahum, Rabbia, Aun Irtaza, Marriam Nawaz, Tahira Nazir, Momina Masood, Sarang Shaikh, and Emad Abouel Nasr. "A robust framework to generate surveillance video summaries using combination of zernike moments and r-transform and deep neural network." Spinger Multimedia Tools and Applications 82, no. 9 (2023): 13811-13835.

[68] Lin, Dongyun, Yiqun Li, Shudong Xie, Tin Lay Nwe, and Sheng Dong. "Ddr-id: Dual deep reconstruction networks based image decomposition for anomaly detection." Spinger Journal of Ambient Intelligence and Humanized Computing 14, no. 3 (2023): 2125-2139.

[69] Mahmood, Farhan, Daehyeon Jeong, and Jeha Ryu. "A New Approach to Traffic Accident Anticipation With Geometric Features for Better Generalizability." IEEE Access 11 (2023): 29263-29274.

[70] Mahmood, Farhan, Daehyeon Jeong, and Jeha Ryu. "A New Approach to Traffic AccidentAnticipationWithGeometricFeaturesforBetterGeneralizability." IEEEAccess 11(2023):29263-29274.

[71] Taghinezhad, N. and Yazdi, M., 2023. A new unsupervised video anomaly detection using multi-scale feature memorization and multipath temporal information prediction. IEEE Access, 11, pp.9295-9310.

[72] Xing, Peng, and Zechao Li. "Visual anomaly detection via partition memory bank module and error estimation." IEEE Transactions on Circuits and Systems for Video Technology (2023).

[73] Lundström, Adam, Mattias O'nils, Faisal Z. Qureshi, and Axel Jantsch. "Improving deep learning based anomaly detection on multivariate time series through separated anomaly scoring." IEEE Access 10 (2022): 108194-108204.

[74] Barbalau, Antonio, Radu Tudor Ionescu, Mariana-Iuliana Georgescu, Jacob Dueholm, Bharathkumar Ramachandra, Kamal Nasrollahi, Fahad Shahbaz Khan, Thomas B. Moeslund, and Mubarak Shah. "SSMTL++: Revisiting self-supervised multi-task learning for video anomaly detection." Computer Vision and Image Understanding 229 (2023): 103656.

[75] Jezequel, Loic, Ngoc-Son Vu, Jean Beaudet, and Aymeric Histace. "Efficient anomaly detection using self-supervised multi-cue tasks." IEEE Transactions on Image Processing 32 (2023): 807-821.

[76] Wu, Peng, Jing Liu, and Fang Shen. "A deep one-class neural network for anomalous event detection in complex scenes." IEEE Transactions on neural networks and learning systems 31, no. 7 (2019): 2609-2622.

[77] Vosta, S. and Yow, K.C., 2023. KianNet: A violence detection model using an attention-based CNN-LSTM structure. IEEE Access.

[78] Song, J.F., Zhao, H.L., Wen, D.Y. and Xu, X.Y., 2022. Video Anomaly Detection Based on Optical Flow Feature Enhanced Spatio-Temporal Feature Network FusionNet-LSTM-G. IEEE Access, 10, pp.130314-130325.

[79] UCSD Anomaly Detection Dataset http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm