https://doi.org/10.48047/AFJBS.6.Si4.2024.3554-3564

African Journal of Biological Sciences

Literature Review: Violence Detection InVideo Surveillance Systems

Vaishali M Bagade[1*], Dr Jagdish B Helonde[2]

[1*]Research Scholar: Dept Of Electrical And Eletronics Engineering, School Of Engineering And Technology Sandip University Nashik,Vaishalibagadejan22@Gmail.Com,
[2]Supervisior : Dept Of Electrical And Eletronics Engineering, School Of Engineering And Technology Sandip University Nashik, Jbhelonde60@Gmail.Com

*Corresponding Author: Vaishali M Bagade
*Research Scholar: Dept Of Electrical And Eletronics Engineering, School Of Engineering And Technology Sandip University Nashik,Vaishalibagadejan22@Gmail.Com,

Article History
Volume 6, Issue Si4, 2024
Received: 15 April 2024.
Accepted: 15 July 2024
Published: 23 July 2024
doi: 10.48047/AFJBS.6.Si4.2024.3554-3564

Abstract
Violence detection in video surveillance systems is a critical task that has garnered significant attention across research and industry domains.
This literature review explores various approaches to violence detection in video surveillance systems, encompassing traditional methods using handcrafted features and advanced deep learning models. This review provides a comprehensive exploration of diverse methodologies employed in this field, ranging from fundamental approaches leveraging handcrafted features to cutting-edge deep learning models. The review covers motion-based detection methods, deep learning-based approaches, hybrid models, comparative analyses of methodologies, and real-world applications and datasets. By examining these diverse techniques, the study aims to highlight the strengths, weaknesses, and future directions in violence detection research. The review categorizes these methodologies into motion-based detection methods, deep learning-based approaches, hybrid models combining multiple techniques, comparative analyses of methodologies, and practical applications using real-world datasets.

Keywords: Violence detection, video surveillance, motion-based detection, deep learning, hybrid models, comparative analysis, real-world applications

Introduction

Violence detection in video surveillance systems is a critical task in ensuring public safety and security. Traditional methods using handcrafted features have paved the way for more sophisticated techniques involving deep learning models. This review aims to provide a comprehensive overview of the current state of violence detection research, categorizing different methodologies, identifying research gaps, and suggesting future research directions.

Scope of Studies:

The scope of this review includes a wide range of studies on violence detection, from traditional approaches to modern deep learning techniques. Specifically, the review covers:

- **Motion-based detection methods**: Techniques that rely on analyzing movement patterns in video frames to identify violent actions.
- **Deep learning-based approaches**: Methods leveraging neural networks to automatically learn features from large datasets.
- **Hybrid models**: Approaches that combine different techniques to enhance the robustness and accuracy of violence detection.
- **Comparative analysis**: Evaluation of various methodologies to understand their strengths, weaknesses, and suitability for different scenarios.
- **Real-world applications and datasets**: Practical implementations of violence detection systems in real-world settings and the use of benchmark datasets for standardized testing.

Classification of Methodologies

Motion-based Detection Methods

Motion-based detection methods analyze the movement within video frames to detect violence. These methods often rely on tracking the motion patterns and changes to identify potentially violent actions.

1. MoBSIFT and Movement Filtering Algorithm (Febin et al., 2020) [2]

- **Description**: This method combines the Scale-Invariant Feature Transform (SIFT) with a movement filtering algorithm to detect violent actions in video sequences. SIFT features are robust descriptors used to identify and track key points in images, making them effective for motion analysis. The movement filtering algorithm further refines the detection by focusing on significant motion patterns indicative of violence.
- **Advantages**: Robust to variations in scale and rotation; effective in detecting significant movements.
- **Disadvantages**: May struggle with complex scenes and require substantial computational resources for real-time applications.

2. Space-Time Interest Points (STIP) (Laptev, 2005) [14]

- **Description**: STIP identifies significant changes in motion across both spatial and temporal dimensions. This method extracts interest points where there is substantial movement, capturing the essence of dynamic actions within a video.
- **Advantages**: Capable of capturing both spatial and temporal aspects of motion; effective for various action recognition tasks.
- **Disadvantages**: Computationally intensive; may not perform well in low-resolution videos.

3. Improved Trajectories (Wang and Schmid, 2013) [15]

- **Description**: This method enhances trajectory-based features by using dense trajectories and motion boundary histograms. Improved Trajectories track dense points within the video frames and analyze their movement patterns, making it possible to identify complex actions like violence.
- **Advantages**: High accuracy in capturing motion details; robust to background clutter.
- **Disadvantages**: High computational complexity; may require significant preprocessing.

Deep Learning-Based Approaches

Deep learning approaches leverage neural networks to automatically learn features from video data, providing significant advancements over traditional methods.

1. **3D Convolutional Neural Networks (3D CNNs)**
   - **Ullah et al. (2019) [13] and Song et al. (2019) [12]**
     - **Description**: 3D CNNs extend traditional 2D convolutional networks to the temporal dimension, allowing them to extract spatiotemporal features from video data. These models can capture motion and appearance features simultaneously, making them highly effective for violence detection.
     - **Advantages**: High accuracy; can capture complex spatiotemporal patterns.
     - **Disadvantages**: Requires large amounts of data; computationally intensive.

2. **Convolutional Long Short-Term Memory (ConvLSTM)**
   - **Sudhakaran and Lanz (2017) [9]**
     - **Description**: ConvLSTM combines convolutional layers with Long Short-Term Memory (LSTM) networks, enabling the capture of both spatial features and temporal dependencies. This hybrid architecture is particularly effective for sequential data like videos.
     - **Advantages**: Captures long-range dependencies; effective for sequential data.
     - **Disadvantages**: Computationally demanding; may require fine-tuning for optimal performance.

3. **CNN-BiLSTM Models**
   - **Halder and Chatterjee (2020) [11]**
     - **Description**: This approach uses Convolutional Neural Networks (CNNs) to extract spatial features and Bidirectional Long Short-Term Memory (BiLSTM) networks to understand temporal dependencies. The bidirectional nature allows the model to consider information from both past and future frames.
     - **Advantages**: Improved temporal understanding; balances spatial and temporal feature extraction.
     - **Disadvantages**: Higher computational requirements; complexity in model training.

Hybrid Models Combining Multiple Techniques

Hybrid models aim to leverage the strengths of various techniques to enhance the robustness and accuracy of violence detection systems.

1. **AI-assisted Edge Vision (Ullah et al., 2021) [4]**
   - **Description**: This system integrates AI with edge computing for real-time violence detection in IoT-based surveillance networks. The approach combines advanced AI algorithms with edge devices, enabling efficient and timely analysis of video data.
   - **Advantages**: Real-time processing; suitable for distributed surveillance networks.
   - **Disadvantages**: Dependence on edge device capabilities; potential network latency issues.

2. **4D Convolutional Neural Networks (Magdy et al., 2023) [5]**
   - **Description**: The 4D CNN extends traditional CNNs by adding an extra dimension to capture comprehensive spatiotemporal features. This approach aims to improve the detection of complex actions in videos.
   - **Advantages**: Enhanced feature extraction; captures intricate spatiotemporal patterns.
   - **Disadvantages**: Extremely computationally intensive; requires significant memory.

3. **Tuna Swarm Algorithm with Deep Learning (Aldehim et al., 2023) [7]**

- **Description**: This method combines the Tuna Swarm Algorithm, an optimization algorithm, with deep learning models to enhance violence detection. The optimization algorithm fine-tunes the deep learning model, improving its performance.
- **Advantages**: Improved detection accuracy; efficient model optimization.
- **Disadvantages**: Complexity in algorithm integration; computationally demanding.

Comparative Analysis of Methodologies

Performance Metrics

Performance metrics are critical for evaluating the effectiveness of violence detection systems. Common metrics include:

1. **Accuracy**: Measures the overall correctness of the detection system by comparing the number of correct predictions to the total number of predictions.
   - **Importance**: High accuracy indicates a reliable system that can correctly identify violent actions most of the time.
2. **Precision and Recall**:
   - **Precision**: The ratio of correctly identified violent actions to the total number of actions identified as violent. High precision indicates fewer false positives.
   - **Recall**: The ratio of correctly identified violent actions to the total number of actual violent actions. High recall indicates fewer false negatives.
   - **Importance**: Both metrics are crucial for balancing the trade-off between false positives and false negatives.

3. **F1-Score**: The harmonic mean of precision and recall, providing a single metric that balances both.
   - **Importance**: Useful when there is an uneven class distribution or when both precision and recall are equally important.

4. **Processing Time**: Measures the efficiency of the detection method in terms of time required to process video data.
   - **Importance**: Critical for real-time applications where quick detection is necessary to respond to violent actions promptly.

Scalability and Generalization

1. **Scalability**: The ability of methods to handle larger datasets and more complex scenarios.
   - **Importance**: Essential for practical applications where surveillance systems must process vast amounts of data continuously.
2. **Generalization**: The robustness of methods to variations in environmental conditions, such as different lighting, camera angles, and background activities.
   - **Importance**: Ensures the detection system performs well across various real-world scenarios and conditions.

Computational Complexity

1. **Memory Usage**: The amount of memory required to process video data, which impacts the feasibility of deploying the method on different hardware platforms.
   - **Importance**: Lower memory usage is advantageous for real-time applications and edge devices with limited resources.

2. Processing Power: The computational resources needed to perform violence detection in real-time.

o **Importance**: Efficient methods with lower processing power requirements are more suitable for deployment in practical surveillance systems, especially in resource-constrained environments.

By examining these methodologies and performance metrics, the review aims to provide a comprehensive understanding of the current state of violence detection research, highlighting the strengths and weaknesses of different approaches and guiding future advancements in this critical field.[40]

Challenges

1. Data Diversity

o **Description**: Video data in surveillance systems exhibits significant variability due to factors such as different camera angles, lighting conditions, and varying background activities.

o **Impact**: This diversity poses challenges for violence detection algorithms, as they must be robust enough to recognize violent actions across different environments and under various conditions.

o **Solution Approaches**:

▪ **Data Augmentation**: Employ techniques to artificially increase the diversity of training data, simulating different lighting conditions, camera viewpoints, and backgrounds.

▪ **Transfer Learning**: Adapt models trained on one dataset to perform effectively on new, diverse datasets by leveraging domain adaptation techniques.

2. False Positives/Negatives

o **Description**: Achieving a balance between precision (minimizing false positives) and recall (minimizing false negatives) is critical for effective violence detection.

o **Impact**: High false positive rates can lead to unnecessary alerts and strain on monitoring resources, while high false negative rates compromise the system's ability to detect actual violent incidents.

o **Solution Approaches**:

▪ **Threshold Adjustment**: Fine-tune detection thresholds to optimize the trade-off between precision and recall based on specific application requirements.

▪ **Advanced Algorithms**: Implement advanced machine learning algorithms that can learn nuanced patterns of violent behavior to reduce false alarms.

▪ **Ensemble Methods**: Combine multiple detectors or classifiers to improve overall detection accuracy and reliability.

3. Scalability

o **Description**: As video surveillance systems grow, the ability of violence detection methods to scale efficiently with increasing data sizes becomes crucial.

o **Impact**: Scalability issues can lead to performance degradation, longer processing times, and increased computational resource requirements.

o **Solution Approaches**:

▪ **Distributed Computing**: Implement distributed systems where computational tasks are distributed across multiple nodes or GPUs to handle large volumes of video data.

▪ **Efficient Algorithms**: Develop algorithms that are optimized for parallel processing and can efficiently utilize hardware resources.

▪ **Edge Computing**: Utilize edge computing to perform initial processing and filtering of video data closer to the source, reducing the load on centralized servers.

4. **Integration with Existing Systems**

o **Description**: Integrating new violence detection methods seamlessly into existing surveillance infrastructure poses integration challenges.

o **Impact**: Compatibility issues and operational disruptions may arise when deploying new detection technologies alongside legacy systems.

o **Solution Approaches**:

▪ **APIs and Standards**: Design detection systems with well-defined APIs and adhere to industry standards to facilitate interoperability with existing surveillance platforms.

▪ **Modular Design**: Adopt a modular approach to system design, allowing components to be easily replaced or upgraded without significant disruption.

▪ **Testing and Validation**: Conduct rigorous testing and validation to ensure the new detection methods integrate smoothly and effectively with existing surveillance operations.

Addressing these challenges requires a combination of technological innovation, robust algorithm development, and careful consideration of real-world deployment scenarios. By tackling these issues, researchers and developers can advance the capabilities of violence detection systems, making them more reliable, efficient, and adaptable to diverse surveillance environments.

Research Gaps and Future Directions

Research Gaps

Despite significant advancements in violence detection, several research gaps and challenges remain:

1. **Limited Generalization**

o **Current Situation**: Many models exhibit strong performance on specific datasets but struggle when applied to different environments or new datasets. This lack of generalization limits the practical applicability of these models in diverse real-world scenarios.

o **Challenges**:

▪ Variations in lighting, camera angles, and background activities across different environments.

▪ Different cultural and contextual interpretations of violent actions.

▪ Limited training data representing the wide variety of real-world conditions.

2. **High Computational Demand**

o **Current Situation**: Deep learning models, particularly those involving 3D CNNs and ConvLSTMs, require substantial computational resources, including powerful GPUs and significant memory.

o **Challenges**:

▪ High computational cost limits the deployment of these models in real-time surveillance systems, especially in resource-constrained environments.

▪ Increased power consumption and operational costs.

▪ Need for specialized hardware, which may not be feasible for widespread deployment.

3. **Lack of Standardized Benchmarks**

o **Current Situation**: The inconsistent use of benchmark datasets across studies hinders the comparability of different methods. Researchers often use custom datasets or modify existing ones, making it difficult to assess the true performance of various approaches.

o **Challenges**:

- Difficulty in comparing and reproducing results across studies.
- Lack of standardized evaluation metrics and protocols.
- Fragmentation in the research community, slowing down the overall progress.

4. Real-Time Challenges

o **Current Situation**: Ensuring low-latency processing for real-time surveillance applications remains a challenge due to the high computational demand and the need for quick decision-making.

o **Challenges**:

- Balancing the trade-off between accuracy and processing speed.
- Managing data throughput and latency in real-time systems.
- Integrating detection systems with existing surveillance infrastructure without compromising performance.

Future Research Directions

Future research should focus on addressing these gaps and overcoming the challenges to advance the field of violence detection:

1. Improved Generalization

o **Approaches**:

- **Transfer Learning**: Utilize transfer learning techniques to adapt pre-trained models to new datasets and environments.
- **Data Augmentation**: Implement extensive data augmentation strategies to simulate a variety of real-world conditions and improve model robustness.
- **Domain Adaptation**: Develop domain adaptation techniques to minimize the performance gap between the training and target domains.

2. Efficient Algorithms

o **Approaches**:

- **Model Compression**: Employ model compression techniques such as pruning, quantization, and knowledge distillation to reduce the size and computational complexity of deep learning models.
- **Lightweight Architectures**: Design lightweight neural network architectures tailored for low-power devices and real-time processing.
- **Optimization Algorithms**: Explore advanced optimization algorithms to speed up training and inference without sacrificing accuracy.

3. Standardized Datasets

o **Approaches**:

- **Benchmark Creation**: Establish standardized benchmark datasets that encompass a wide range of scenarios, environments, and types of violent actions.
- **Evaluation Protocols**: Develop and adopt uniform evaluation protocols and metrics to ensure consistent assessment of different methods.
- **Open Access**: Promote open access to datasets and encourage the research community to contribute to and use these benchmarks.

4. Edge Computing

o **Approaches**:

▪ **Edge AI**: Leverage edge computing to process video data closer to the source, reducing latency and bandwidth requirements.

▪ **Distributed Systems**: Implement distributed surveillance systems where computational tasks are shared between edge devices and central servers.

▪ **Adaptive Models**: Design adaptive models that can dynamically adjust their complexity based on the available computational resources.

5. Multimodal Approaches

o **Approaches**:

▪ **Sensor Fusion**: Integrate data from multiple sensors, such as audio, text, and video, to enhance the understanding of violent actions.

▪ **Multimodal Learning**: Develop multimodal learning frameworks that can process and combine different types of data, improving the overall accuracy and robustness of violence detection systems.

▪ **Contextual Analysis**: Incorporate contextual information, such as location and time, to provide a more comprehensive analysis of the detected violent actions.

By focusing on these future research directions, the field of violence detection can achieve significant advancements, leading to more robust, efficient, and practical solutions for real-world applications.

Conclusion

This review highlights the diversity of approaches in violence detection, ranging from traditional handcrafted feature methods to advanced deep learning models.

The literature on violence detection in intelligent surveillance systems reveals significant advancements in methodologies and applications. While deep learning and hybrid models show promise in improving detection accuracy, challenges related to real-time processing and While significant progress has been made, challenges remain in generalization, computational efficiency, and real-time applicability. Addressing these challenges through advanced technologies and comprehensive datasets will be crucial for the future development of reliable and efficient violence detection systems. Future research should focus on addressing these gaps, improving model robustness, and leveraging emerging technologies like edge computing for enhanced violence detection in video surveillance systems.

References

1. Ullah, F.U.M., Obaidat, M.S., Muhammad, K., Ullah, A., Baik, S.W., Cuzzolin, F., Rodrigues, J.J. and de Albuquerque, V.H.C., 2022. An intelligent system for complex violence pattern analysis and detection. International Journal of Intelligent Systems, 37(12), pp.10400-10422.
2. Febin, I.P., Jayasree, K. and Joy, P.T., 2020. Violence detection in videos for an intelligent surveillance system using MoBSIFT and movement filtering algorithm. Pattern Analysis and Applications, 23(2), pp.611-623.
3. Asad, M., Yang, J., He, J., Shamsolmoali, P. and He, X., 2021. Multi-frame feature-fusion-based model for violence detection. The Visual Computer, 37, pp.1415-1431.
4. Ullah, F.U.M., Muhammad, K., Haq, I.U., Khan, N., Heidari, A.A., Baik, S.W. and de Albuquerque, V.H.C., 2021. AI-assisted edge vision for violence detection in IoT-based

industrial surveillance networks. IEEE Transactions on Industrial Informatics, 18(8), pp.5359-5370.

5. Magdy, M., Fakhr, M.W., Maghraby, F.A.: Violence 4D: violence detection in surveillance using 4D convolutional neural networks. IET Comput. Vis. 17(3), 282–294 (2023).
6. Kwar, K., Attar, V.: Deep learning model based on cascaded autoencoders and one-class learning for detection and localization of anomalies from surveillance videos. IET Biome. 11(4), 289–303 (2022).
7. Aldehim, G., Asiri, M.M., Aljebreen, M., Mohamed, A., Assiri, M., Ibrahim, S.S.: Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems. IEEE Access 11, 95104–95113 (2023).
8. Kwan-Loo, K.B., Ortíz-Bayliss, J.C., Conant-Pablos, S.E., Terashima-Marín, H., Rad, P.: Detection of Violent Behavior Using Neural Networks and Pose Estimation. IEEE Access 10, 86339-86352 (2022).
9. Sudhakaran, S., Lanz, O.: Learning to detect violent videos using convolutional long short-term memory. 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-6, 2017.
10. Vijeikis, R., Raudonis, V., Dervinis, G.: Efficient violence detection in surveillance. Sensors 22(6), 2216 (2022).
11. Halder, R., Chatterjee, R.: CNN-BiLSTM model for violence detection in smart surveillance. SN Computer Science 1(4), 201 (2020).
12. Song, W., Zhang, D., Zhao, X., Yu, J., Zheng, R., Wang, A.: A novel violent video detection scheme based on modified 3D convolutional neural networks. IEEE Access 7, 39172-39179 (2019).
13. Ullah, F.U.M., Ullah, A., Muhammad, K., Haq, I.U., Baik, S.W.: Violence detection using spatiotemporal features with 3D convolutional neural network. Sensors 19(11), 2472 (2019).
14. Laptev, I.: On space-time interest points. International Journal of Computer Vision 64, 107-123 (2005).
15. Wang, H., Schmid, C.: Action recognition with improved trajectories. IEEE International Conference on Computer Vision, pp. 3551-3558, 2013.
16. Xu, L., Gong, C., Yang, J., Wu, Q., Yao, L.: Violent video detection based on MoSIFT feature and sparse coding. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3538-3542, 2014.
17. Chen, D., Wactlar, H., Chen, M.Y., Gao, C., Bharucha, A., Hauptmann, A.: Recognition of aggressive human behavior using binary local motion descriptors. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5238-5241, 2008.
18. Neshige, R., Kuroda, Y., Kakigi, R., Fujiyama, F., Matoba, R., Yarita, M., Lüders, H., Shibasaki, H.: Event-related brain potentials as indicators of visual recognition and detection of criminals by their use. Forensic Science International 51(1), 95-103 (1991).
19. Kacprzyk, J.: Advances in intelligent systems and computing. Springer, 2012.
20. Muhammad, K., Khan, S., Elhoseny, M., Ahmed, S.H., Baik, S.W.: Efficient fire detection for uncertain surveillance environment. IEEE Transactions on Industrial Informatics 15(5), 3113-3122 (2019).
21. Muhammad, K., Hussain, T., Baik, S.W.: Efficient CNN based summarization of surveillance videos for resource-constrained devices. Pattern Recognition Letters 130, 370-375 (2020).
22. Mabrouk, A.B., Zagrouba, E.: Spatio-temporal feature using optical flow based distribution for violence detection. Pattern Recognition Letters 92, 62-67 (2017).

23. Accattoli, S., Sernani, P., Falcionelli, N., Mekuria, D.N., Dragoni, A.F.: Violence detection in videos by combining 3D convolutional neural networks and support vector machines. Applied Artificial Intelligence 34(4), 329–344 (2020).
24. Fradi, H., Luvison, B., Pham, Q.C.: Crowd behavior analysis using local mid-level visual descriptors. IEEE Transactions on Circuits and Systems for Video Technology 27(3), 589–602 (2016).
25. Al-Dhamari, A., Sudirman, R., Mahmood, N.H.: Transfer deep learning along with binary support vector machine for abnormal behavior detection. IEEE Access 8, 61085–61095 (2020).
26. Sayed, G.I., Soliman, M.M., Hassanien, A.E.: A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. Computers in Biology and Medicine 136, 104712 (2021).
27. Mohammadi-Balani, A., Nayeri, M.D., Azar, A., Taghizadeh-Yazdi, M.: Golden eagle optimizer: A nature-inspired metaheuristic algorithm. Computers & Industrial Engineering 152, 107050 (2021).
28. DCSASS dataset: https://www.kaggle.com/datasets/mateohervas/dcsass-dataset accessed on February 2024.
29. Huang, C., Wang, H.: A novel key-frames selection framework for comprehensive video summarization. IEEE Transactions on Circuits and Systems for Video Technology 30(2), 577-589 (2019).
30. Huang, Z., Zhao, H., Zhan, J., Li, H.: A multivariate intersection over union of SiamRPN network for visual tracking. The Visual Computer 38(8), 2739–2750 (2022).
31. Gordon, D., Farhadi, A., Fox, D.: Re^3: Real-time recurrent regression networks for visual tracking of generic objects. IEEE Robotics and Automation Letters 3(2), 788–795 (2018).
32. Held, D., Thrun, S., Savarese, S.: Learning to track at 100 fps with deep regression networks. 14th European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 2016.
33. Ahmed, M., Ramzan, M., Khan, H.U., Iqbal, S., Khan, M.A., Choi, J.I., Nam, Y., Kadry, S.: Real-time violent action recognition using key frames extraction and deep learning.
34. Colque, R.V.H.M., Caetano, C., de Andrade, M.T.L., Schwartz, W.R.: Histograms of optical flow orientation and magnitude and entropy to detect anomalous events in videos. IEEE Transactions on Circuits and Systems for Video Technology 27(3), 673–682 (2016).
35. Rassem, T.H., Khoo, B.E.: Completed local ternary pattern for rotation invariant texture classification. The Scientific World Journal, 2014.
36. Jabid, T., Kabir, M.H., Chae, O.: Local directional pattern (LDP)–A robust image descriptor for object recognition. 7th IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 482–487, 2010.
37. Jabid, T., Kabir, M.H., Chae, O.: Gender classification using local directional pattern (LDP). 20th International Conference on Pattern Recognition, pp. 2162–2165, 2010.
38. Xue, W., Zhang, L., Mou, X., Bovik, A.C.: Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. IEEE Transactions on Image Processing 23(2), 684-695 (2013).
39. Kalshetty, R., Parveen, A.: Abnormal event detection model using an improved ResNet101 in context aware surveillance system. Cognitive Computation and Systems 5(2), 153–167 (2023).
40. Siddique, L.A., Junhai, R., Reza, T., Khan, S.S., Rahman, T.: Analysis of Real-Time Hostile Activity Detection from Spatiotemporal Features Using Time Distributed Deep CNNs, RNNs and Attention-Based Mechanisms. IEEE 5th International Conference on Image Processing Applications and Systems (IPAS), pp. 1-6, 2022.

41. Baba, M., Gui, V., Cernazanu, C., Pescaru, D.: A sensor network approach for violence detection in smart cities using deep learning. Sensors 19(7), 1676 (2019).
42. Honarjoo, N., Abdari, A., Mansouri, A.: Violence detection using one-dimensional convolutional networks. 12th International Conference on Information and Knowledge Technology (IKT), pp. 188-191, 2021.
43. Mehmood, A.: Efficient anomaly detection in crowd videos using pre-trained 2D convolutional neural networks. IEEE Access 9, 138283-138295 (2021).
44. Wang, Z., Yan, W., Li, Z., Huang, M., Fan, Q., Wang, X.: Domestic Violence Crisis Recognition Method based on Bi-LSTM+ Attention. 8th Annual International Conference on Network and Information Systems for Computers (ICNISC), pp. 569-575, 2022.