

A Review of Road Surveillance using Artificial Intelligence

Abhijith Chandran^{*}, Ajay Janardhanan[#], Akhil Siji Chandy[&], Amaljith KK[@], Jyothis Joseph⁺
Department of Computer Science and Engineering, College of Engineering Kidangoor, Kottayam, Kerala, India

^{*}abhijithchandran3000@gmail.com, [#]ajayjanardhanan30@gmail.com, [&]akhilsiji7@gmail.com,
[@]amaljitkk09@gmail.com,
⁺jyothis@ce-kgr.org

Abstract—Road surveillance has become highly important in the modern day due to the rise in crime and accident rates. Traditional manual monitoring often limits scalability, slows decision-making, and increases operational costs. There are many studies being carried out, attempting to solve these issues. Although it is difficult to keep up with the increasing number of vehicles on the roads, technology provides ways to improve surveillance systems through the use of machine learning to detect anomalies such as road accidents or fights. This survey paper examines advancements in anomaly detection using machine learning. Additionally, it highlights the limitations in current anomaly detection systems and drawbacks of machine learning models that are trained on real world footage and identifies areas for future improvement, paving the way for more robust ways to detect anomalies such as accidents and fights.

Index Terms—Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (Bi-LSTM), Computer Vision (CV), Explainable AI (XAI).

I. INTRODUCTION

In today's world, urban environments are increasingly dependent on surveillance systems to ensure public safety and maintain law and order. However, traditional CCTV systems rely mainly on manual monitoring, making it difficult to efficiently analyze large amounts of visual data and detect critical events in real time. This limitation often delays response actions in cases such as thefts, accidents, or missing person incidents, where every second is crucial. The growing availability of high-resolution cameras, coupled with advancements in Artificial Intelligence (AI) and Computer Vision, presents an opportunity to overcome these challenges through intelligent automation.

This survey focuses on reviewing recent advancements in intelligent surveillance systems, particularly those aimed at identifying missing persons and stolen vehicles using AI-driven techniques. It explores the evolution of traditional monitoring methods toward automated, data-driven systems that incorporate face recognition, vehicle detection, anomaly detection, and route reconstruction. The review also highlights the datasets, algorithms, and evaluation metrics employed in previous works to provide a comprehensive understanding of the current research landscape. By analyzing various approaches and identifying their strengths and limitations, this study lays the groundwork for developing a robust, real-time CCTV surveillance solution that can improve public safety, accelerating search operations, and supporting law enforcement with evidence-based decision-making.

II. LITERATURE SURVEY

A. Attire-Based Anomaly Detection in Restricted Areas Using YOLOv8 for Enhanced CCTV Security

Abdul Aziz A.B et al. [1] explores an innovative security enhancement approach, employing advanced image analysis and soft computing. The focus is on an intelligent surveillance system that detects unauthorized individuals in restricted areas by analyzing attire. This research harnesses the cutting-edge capacities of advanced deep learning technology, prominently featuring YOLOv8 (You Only Look Once version 8)—a preeminent object detection algorithm.

The methodology involves curating a diverse dataset encompassing various attire patterns associated with authorized personnel in restricted areas, employing manual annotation to label images with corresponding attire categories to ensure the model's accuracy during training. Then, the YOLOv8 model is trained and fuzzy logic is employed to enhance the system's adaptability to dynamic environments, variations in lighting etc. The dataset is divided into training, validation, and testing sets to evaluate the model's performance objectively by employing k-fold cross-validation. Finally, the trained YOLOv8 model is integrated into a real-time surveillance system for seamless anomaly detection after addressing hardware requirements and ensuring compatibility with CCTV infrastructure.

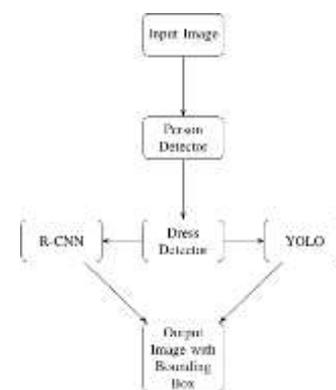


Fig. 1. Simplified workflow architecture of the proposed system [1]

As shown in Fig.1, the process initiates with the input image, which undergoes preprocessing. The preprocessed image

is then fed into a Backbone Network, a crucial element which captures hierarchical features, enabling the subsequent detectors to make informed decisions. The first specialized detector focuses on identifying individuals within the image, utilizing the features extracted by the Backbone Network to localize and classify persons. Simultaneously, the Mask Detector operates on the same set of features, aiming to identify the presence of facial masks on the detected individuals. The Dress Detector focuses on identifying clothing items within the image. The Mask Detector and Dress Detector share information with the YOLO object detection algorithm, which predicts bounding boxes and class probabilities for multiple objects in real-time. Running in parallel to YOLO, the Mask R-CNN (Region-based Convolutional Neural Network) specializes in mask segmentation. The outputs from YOLO and Mask R-CNN undergo a critical stage of Feature Fusion, which integrates information from both detectors, enriching the analysis with combined spatial and semantic information. The final stage of the pipeline produces an output image enriched with bounding boxes outlining detected persons and dresses, along with accurately segmented masks.

The successful implementation of this system in real-world scenarios could significantly elevate security standards and mitigate potential threats, marking a pivotal step forward in the realm of advanced surveillance technology.

B. Analysis of Real-Time Hostile Activity Detection from Spatiotemporal Features Using Time Distributed Deep CNNs, RNNs and Attention-Based Mechanisms

Labib Ahmed Siddique et al. [2] explores deep learning video classification techniques to detect violence as they are happening. Although CCTV cameras help a lot in increasing security, these systems require constant human interaction and monitoring. To eradicate this issue, the proposed deep learning models that can be easily integrated with IOT such as LRCN (both custom CNN and pre-trained), C3D, ConvLSTM, and CNN-Transformers to help automate surveillance systems. The primary objectives of this research include creating deep learning models to detect violent or non-violent human behavior in an automated way, processing raw CCTV footage to extract features.

In the data set used, there are 350 clips that are in MP4 format (H.24 codec) where 120 clips contain non-violent videos and 230 clips contain violent videos. The frame rate per video is 30 fps, and the resolution is fixed to 1920x1080 pixels for all the clips. The idea of Long-term Recurrent Convolutional Network is to extract spatial features from each frame using convolution neural networks. The outputs of these convolutional networks are passes into a Bi-LSTM (Bidirectional Long Short-Term Memory) network, which fuses temporal features onto extracted spatial features to classify them.

Amongst all the models, the CNN-Transformer model performs the most optimally with 0.74 being the F1 score for non violent and 0.79 being for the violent case.

This implementation of this system provides us with models that can help prevent crime by tracking and analyzing the

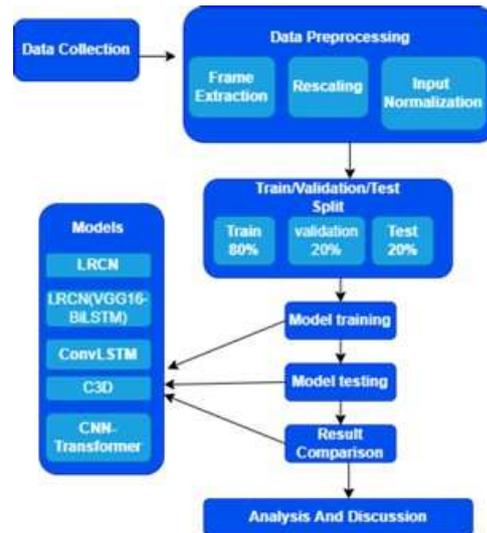


Fig. 2. The flow chart of the proposed hostile behavior detection model [2]

footage from CCTV cameras in real time, making the job of the investigating team easier and help in prevailing justice for the victims.

C. Enhancing Surveillance Anomaly Detection with Keyframes and Explainable Inception Model

Muhammad Salman et al. [3] present a sophisticated approach to improving the accuracy and interpretability of intelligent surveillance systems through two proposed models — SilentFrm and XAI-Inv3. The study addresses the limitations of traditional surveillance systems, which often fail in real-time anomaly detection due to complex environments, lighting variations, and diverse human appearances.

The SilentFrm framework enhances anomaly detection efficiency by selecting keyframes — representative frames that reduce redundancy in video sequences. It leverages frame difference analysis and adaptive thresholding to identify frames that capture critical events, thereby reducing computational overhead while maintaining precision. This process ensures faster anomaly identification and scalability across large-scale deployments.

Complementing SilentFrm, XAI-Inv3 (Explainable InceptionV3) integrates explainable artificial intelligence (XAI) techniques such as Grad-CAM and guided backpropagation into the InceptionV3 architecture. This fusion not only achieves high detection accuracy but also provides interpretability, allowing visualization of regions within video frames that influence the model’s decision. Consequently, security personnel can understand and trust the system’s outputs. Fig.3 depicts the overall structure of the system, integrating keyframe extraction, deep learning-based detection, and explainable visualization modules. By combining efficiency with interpretability, the proposed system ensures real-time,

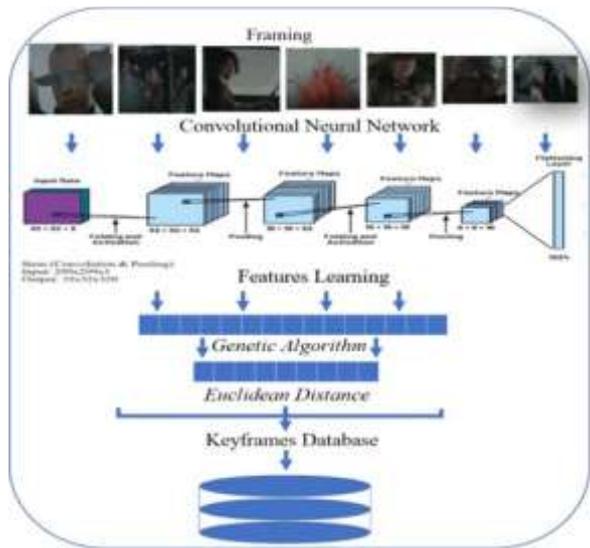


Fig. 3. Proposed keyframes selection framework [3]

scalable, and transparent anomaly detection suitable for modern surveillance environments.

Experiments conducted on multiple benchmark datasets — including Hockey Fight, Violent Flow, Real-Life Violence Situation, UCF-Crime, and ShanghaiTech — demonstrate the proposed model’s superior performance. SilentFrm and XAI-Inv3 achieved up to 99% accuracy in violence detection and 92% accuracy in real-world surveillance datasets, outperforming existing state-of-the-art models.

D. A Lightweight Face Quality Assessment Framework to Improve Face Verification Performance in Real-time Screening Applications

Ahmed Aman Ibrahim et al. [4] introduce a lightweight and efficient face quality assessment (FQA) framework designed to enhance face verification accuracy in real-time surveillance and screening scenarios. Recognizing that low-quality images — often degraded by motion blur, poor lighting, occlusion, or pose variation — severely compromise recognition systems, this study proposes an intelligent pre-processing module that filters poor-quality face inputs before verification.

The framework operates through a three-stage pipeline involving face detection, landmark normalization, and classification. Using RetinaFace for face detection, key facial landmarks (eyes, nose tip, and lip corners) are extracted and normalized relative to the face bounding box to eliminate scale and pose variations. These normalized coordinates are then input into a Random Forest Regression model, which classifies the image as high or low quality. This machine learning-based approach provides a fast, lightweight solution suitable for deployment in constrained surveillance environments.

To validate the model, the authors constructed a real-world CCTV dataset comprising over 600 subjects captured under unconstrained conditions by the Dubai Police. The model achieved an impressive 96.67% accuracy, outperforming state-

of-the-art FQA methods such as SER-FIQ, FaceQNet, and MagFace. The results demonstrate a 99.7% reduction in False Rejection Rate (FRR) and a 15% improvement in the cosine similarity scores when integrated with the ArcFace verification model.

By combining simplicity with strong predictive power, this framework significantly enhances the reliability of face verification systems used in surveillance and forensic applications. It ensures that only high-quality, properly aligned facial images proceed to the recognition stage — a critical improvement for real-time, high-stakes screening systems like law enforcement CCTV networks and border control.

This research highlights how landmark-based quality estimation can act as a lightweight yet powerful preprocessing layer, bridging the gap between traditional biometric quality assessment and modern AI-driven recognition pipelines.

E. Gta-Crime: A Synthetic Dataset and Generation Framework for Fatal Violence Detection with Adversarial Snippet-Level Domain Adaptation

Seongho Kim et al. [5] present GTA-Crime, a novel synthetic dataset and generation framework designed to advance fatal violence detection in surveillance systems through a snippet-level domain adaptation strategy. Recognizing that real-world datasets rarely include critical violent incidents such as shootings or stabbings—mainly due to ethical and safety concerns—the authors utilize the Grand Theft Auto V (GTA5) virtual environment to simulate these rare yet vital events.

The GTA-Crime dataset contains 532 CCTV-style videos, including 270 fatal violence clips (124 stabbings and 146 shootings) and 262 normal clips, captured under varying conditions such as weather, time of day, and camera angle. Each video spans 13 seconds at 1920×1080 resolution, recorded from multiple CCTV perspectives, offering both frame-level and video-level annotations. This rich variability ensures close resemblance to real surveillance footage while maintaining control over scenario parameters like lighting and location.

To ensure that models trained on synthetic data can generalize to real-world settings, the study introduces an Adversarial Snippet-Level Domain Adaptation (ASDA) approach based on Wasserstein adversarial training (WGAN-GP). This method aligns feature distributions between the synthetic GTA-Crime dataset and the real UCF-Crime dataset at the snippet level. Specifically, class-wise adaptation aligns “stabbing” clips with “fighting,” “shooting” with “shooting,” and “normal” with “normal” classes across domains, minimizing domain discrepancies and improving transfer learning performance.

Fig. 5 outlines the workflow of the proposed framework, highlighting the alignment process between GTA and UCF feature spaces and its integration with anomaly detection models.

Experiments using multiple Video Anomaly Detection (VAD) models—such as RTFM, UR-DMU, MGFN, CLIP-TSA, and VadCLIP—demonstrate the framework’s effectiveness. Results show that the GTA-Crime dataset achieves com-

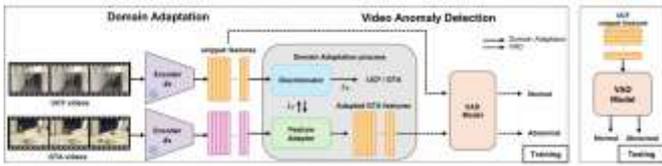


Fig. 4. An overview of the proposed method for improving fatal violence detection [5]

petitive AUC scores (up to 93%) in supervised settings and significantly boosts detection accuracy when domain adaptation is applied. Integrating GTA-Crime features with UCF-Crime improved fatal violence detection performance by over 3% on average compared to training on real data alone.

By combining synthetic data generation with domain adaptation, GTA-Crime provides a scalable, ethical, and versatile resource for developing real-time surveillance systems capable of detecting life-threatening events. This work establishes a significant milestone in video anomaly detection research, addressing the scarcity of fatal incident data while ensuring improved cross-domain generalization for safety-critical AI applications.

F. Application of 2D Homography for High Resolution Traffic Data Collection using CCTV Cameras

Linlin Zhang et al. [6] present a solution to a key challenge in utilizing existing infrastructure-mounted CCTV cameras: the difficulty in extracting high-resolution data (vehicle counts, speed, and acceleration) due to camera distortion, partial occlusion, and complex calibration requirements. The work aims to augment traditional data collection with video analytics to create a robust and high-resolution traffic monitoring system for ITS, which can also power Digital Twin technology.

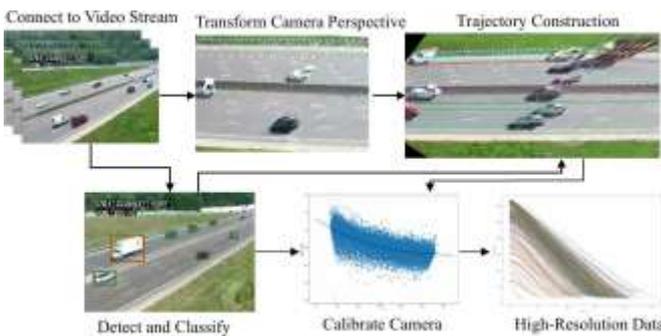


Fig. 5. Proposed system framework [5]

Fig. 6 shows the implemented three-stage video analytics framework, consisting of:

Object Recognition: A state-of-the-art deep learning model (YOLOv5 is mentioned in the methodology) is used to detect and classify vehicles.

Perspective Transformation: A 2D homography technique is applied to transform the distorted CCTV camera perspective into a Bird's-Eye View (BEV). This key step corrects camera

distortion and reduces the effect of partial occlusion, thereby improving data accuracy.

Vehicle Trajectory Reconstruction: The system uses object-tracking algorithms, such as Motpy and BYTETrack, to construct individual vehicle trajectories in the BEV. The system also includes an automatic camera calibration with a two-layer matrix system to convert pixel coordinates into actual real-world measurements, allowing for accurate speed and acceleration extraction.

The framework yielded high accuracy results, showing an error rate of about +/- 4.5% for directional traffic counts and a speed bias of less than 10% MSE when compared to ground truth probe data sources. The ability to extract high-resolution data such as individual vehicle trajectories, speeds, and accelerations has significant implications for improving traffic management and identifying high-risk areas.

G. Finetuning Yolov9 for Vehicle Detection: Deep Learning for Intelligent Transportation Systems in Dhaka, Bangladesh

Shahriar Ahmad Fahim [7] addresses the pressing transportation challenges, such as extreme traffic congestion, in megacities like Dhaka, Bangladesh, by proposing an advancement in Intelligent Transportation Systems (ITS). The study utilizes deep learning, a subset of Artificial Intelligence (AI), for traffic surveillance and to identify mobility patterns.

The paper introduces a fine-tuned object detector, the YOLOv9 model, specifically trained on a Bangladesh-based dataset, Poribohon-BD, which contains 9,058 labeled images of 15 native vehicle types. Fine-tuning is necessary because native vehicles in Dhaka have unique features that are often dissimilar from those in general-purpose international datasets. The model uses the lighter YOLOv9-C architecture, accommodating environments with low computational resources.

The fine-tuned YOLOv9 model achieved a state-of-the-art performance with a mean Average Precision (mAP) of 0.934 at an Intersection over Union (IoU) threshold of 0.5. Beyond detection, the study proposes a conceptual framework for a city-wide vehicle detection system by suggesting the model be deployed on CCTVs and processing the output data into a graph structure.

H. Integrating Generative Adversarial Networks and Convolutional Neural Networks for Enhanced Traffic Accidents Detection and Analysis

Zhenghao Xi et al. [8] present an advanced deep learning framework integrating Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs) to improve the accuracy and efficiency of accident detection using CCTV surveillance footage. The motivation behind this research arises from the growing number of global traffic accidents and the urgent need for automated, real-time surveillance systems that can effectively detect accidents and alert authorities. Traditional monitoring systems often rely on manual observation or supervised learning approaches that demand extensive labeled datasets—an impractical requirement given the rarity of accident data. To address this data scarcity, the

authors employ GANs for synthetic data generation, creating realistic accident and non-accident scenarios to enrich the training dataset, while CNNs are utilized for feature extraction and classification of accident events.

The dataset used consisted of accident and non-accident video frames extracted from publicly available YouTube CCTV footage, which were resized, enhanced, and normalized to improve training efficiency. GAN-based Deep Convolutional GAN (DCGAN) models generated synthetic images to compensate for the lack of diverse accident samples, ensuring balanced data distribution. The system architecture demonstrates a workflow in which the GAN-generated data feeds into the CNN-based classifier, providing a robust mechanism for real-time detection and automated alerting.

The integration of deep learning with intelligent traffic monitoring systems provides a foundation for future smart cities, where real-time accident detection and emergency response are seamlessly automated. This research not only demonstrates the practical implementation of AI in transportation safety but also sets a benchmark for intelligent surveillance systems by achieving both accuracy and scalability. The authors conclude that incorporating synthetic data generation and transformer-based architectures can revolutionize traffic incident management, paving the way for efficient, AI-driven accident prevention and emergency coordination.

I. Anomaly Alert System using CCTV surveillance

Manu Y. M. et al. [9] propose an Anomaly Alert System designed to enhance fire safety through real-time CCTV-based image processing integrated with Raspberry Pi technology. Traditional fire alarm systems that rely on smoke or heat sensors often face challenges such as delayed detection, false alarms, and high installation costs. To overcome these limitations, this study introduces a vision-based fire detection framework that uses existing CCTV infrastructure to monitor environments continuously and identify fire outbreaks at their earliest stage. The system's primary objective is to provide a cost-effective, automated, and accurate fire detection mechanism suitable for industrial and commercial facilities.

The proposed system utilizes the Raspberry Pi module as the central processing unit, interfaced with a Pi Camera that continuously captures video frames from surveillance feeds. The captured images are processed using OpenCV-based algorithms to analyze color intensity and detect fire-like patterns in real-time.

The architecture of the system comprises three major components—image acquisition, fire detection using Haar-Cascade classifiers, and alert generation. The process begins with continuous monitoring of video streams, followed by feature extraction using Histogram of Oriented Gradients (HOG) and Gray Level Co-occurrence Matrix (GLCM) techniques to detect fire boundaries and intensity. Subsequently, Support Vector Machine (SVM) classifiers are applied for accurate categorization of fire events based on extracted visual parameters.

J. Real-Time Anomaly Recognition Through CCTV Using Neural Networks

Virender Singh et al. [10] present a deep learning-based anomaly recognition framework designed to automate real-time surveillance using CCTV video feeds. The growing volume of disruptive and violent activities in public spaces, coupled with the limitations of manual monitoring, necessitates an intelligent system that can identify suspicious or aggressive behavior autonomously. This study proposes a hybrid model combining Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to detect, classify, and alert authorities about abnormal activities from live video streams. The proposed architecture employs a two-stage deep learning pipeline where CNNs extract spatial features from video frames, and RNNs analyze temporal dependencies to interpret sequences of actions over time. The InceptionV3 model is utilized for transfer learning, leveraging pre-trained weights from large-scale datasets like ImageNet to reduce training time and improve generalization. The system can recognize 12 distinct anomalies including Abuse, Burglary, Explosion, Shooting, Fighting, Shoplifting, Road Accidents, Arson, Robbery, Stealing, Assault, and Vandalism, making it suitable for real-world surveillance scenarios.

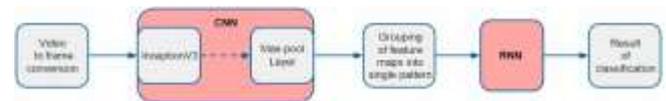


Fig. 6. Workflow of Anomaly Recognition System [10]

III. COMPARATIVE STUDY

Recent advancements in video surveillance and anomaly detection, driven by deep learning, computer vision, and multi-model integration, face challenges such as real-time performance, data quality, and environmental variability. Table I summarizes selected studies, highlighting the primary algorithms and models employed, along with their unique contributions, including violence detection, fire monitoring, vehicle tracking, and explainable AI implementations. By combining algorithmic innovation with practical deployment considerations, these approaches contribute to more robust, adaptable, and intelligent surveillance systems capable of addressing diverse monitoring needs.

Overall, the comparative analysis shows that modern surveillance systems are evolving toward multi-module architectures combining detection, tracking, recognition, and explainability. However, common challenges remain, including illumination variation, camera angle differences, computational cost, and the lack of diverse real-world datasets. Therefore, an effective intelligent surveillance solution should integrate multi-camera analysis, robust preprocessing, and real-time processing capabilities to achieve reliable performance across varied environments.

TABLE I
COMPARATIVE ANALYSIS OF STATE-OF-THE-ART VIDEO AND ANOMALY DETECTION METHODOLOGIES

Sl. No.	Author(s)	Primary Algorithm / Model	Unique Strength
1	Abdul Aziz A.B et al. [1]	YOLOv8 + Mask R-CNN + Fuzzy Logic	Adaptable to varied environments, capable of rule-driven abnormal-dress alerts.
2	Labib Ahmed Siddique et al. [2]	LRCN (CNN + Bi-LSTM), C3D, CNN-Transformer	Temporal modelling of violent activities; Transformer achieves high F1 for both violent and non-violent classes.
3	Muhammad Salman et al. [3]	Key-Frame Selection + InceptionV3 (XAI-Inv3)	Explainability (Grad-CAM) and key-frame reduction, high accuracy with lower computational load.
4	Ahmed Aman Ibrahim et al. [4]	RetinaFace + Landmark Normalization + Random Forest	Face-quality assessment framework to filter low-quality CCTV frames before recognition, improving verification accuracy.
5	Seongho Kim et al. [5]	ASDA (WGAN-GP + Adversarial Snippet Domain Adaptation)	Leveraged synthetic GTA-Crime dataset with adversarial domain adaptation to tackle rare violent events.
6	Linlin Zhang et al. [6]	YOLOv5 + 2D Homography + BYTETrack	Converted CCTV perspectives to Bird's-Eye View (BEV) for precise trajectory and speed estimation across camera views.
7	Shahriar Ahmad Fahim et al. [7]	Fine-Tuned YOLOv9-C	Localized vehicle detection for Bangladeshi roads with high mAP (0.934), demonstrating the value of regional fine-tuning.
8	Zhengkao Xi et al. [8]	DCGAN + CNN / ViT / FTCNN	GAN-based synthetic data generation improved accident detection; Superior contextual understanding.
9	Manu Y. M et al. [9]	OpenCV + Haar Cascade + GLCM + SVM	Low-cost real-time fire anomaly detection system on Raspberry Pi with detection accuracy around 99%.
10	Virender Singh et al. [10]	InceptionV3 + LSTM	Multi-class anomaly recognition (12 categories) on low-quality CCTV videos achieving 97% accuracy.

IV. CONCLUSIONS

Advancements in video surveillance and anomaly detection are enabling smarter, more responsive monitoring systems. Through deep learning, computer vision, and multi-model integration, researchers have developed frameworks capable of detecting a wide range of anomalies—from violence and fire to vehicle tracking—across diverse environments. The reviewed works demonstrate how techniques such as YOLO, Transformer-based models, GANs, and explainable AI improve accuracy, efficiency, and adaptability in real-world CCTV applications. Looking forward, integrating multi-camera coordination, real-time analytics, and robust preprocessing methods will be key to building intelligent surveillance systems that are both reliable and scalable. By combining technological innovation with practical deployment considerations, these approaches pave the way for safer, more secure, and highly automated monitoring solutions.

REFERENCES

[1] Abdul Aziz A.B and Aindri Bajpai, "Attire-based anomaly detection in restricted areas using yolov8 for enhanced cctv security," 2024.

[2] T. S. Labib Ahmed Siddique, Rabita Junhai and Tanvir Rahman, "Analysis of real-time hostile activity detection from spatiotemporal features using time distributed deep cnns, rnns and attention-based mechanisms," 2023.

[3] S. A. F. A. Muhammad Salman, Naveed Abbas and Tanzila Saba, "Enhancing surveillance anomaly detection with keyframes and explainable inception model," *Egyptian Informatics Journal* 31, 2025.

[4] A. A. B. S. I. Ahmed Aman Ibrahim, Hamad Mansour Alawar and Hassan Ugail, "A lightweight face quality assessment framework to improve face verification performance in real-time screening applications," 2025.

[5] H. Seongho Kim, Sejong Ryu and Je Hyeong Hong, "Gta-crime: A synthetic dataset and generation framework for fatal violence detection with adversarial snippet-level domain adaptation," 2025.

[6] A. Linlin Zhang, Xiang Yu and Y. Abdul Rashid Mussah, "Application of 2d homography for high resolution traffic data collection using cctv cameras," 2023.

[7] Shahriar Ahmad Fahim, "Finetuning yolov9 for vehicle detection: Deep learning for intelligent transportation systems in dhaka, bangladesh," 2024.

[8] Y. Y. Zhengkao Xi, Xiang Liu and Yangyu Zheng, "Integrating generative adversarial networks and convolutional neural networks for enhanced traffic accidents detection and analysis," 2025.

[9] R. Manu YM and Shashikala S V, "Anomaly alert system using cctv surveillance," 2022.

[10] S. Virender Singh and Dr. Pooja Gupta, "Real-time anomaly recognition through cctv using neural networks," 2020.