International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Detection of Road Accidents using Computer Vision

Mylapilli Sri Ram Rajesh[1] , Ms.Manisha Das[2]

[1]B. tech Student, GMR Institute of Technology, Rajam, 532127, India.
[2]Assistant Professor, GMR Institute Of Technology, Rajam, 532127, India

ABSTRACT :

Traffic surveillance is crucial for road safety and managing traffic. Accurate detection of accidents from surveillance videos is vital for incident response and congestion control. The survey presents a deep learning-based method for automatically detecting and locating road accidents in traffic surveillance videos. The approach consists of two main stages: object detection and accident classification. Accidents can be detected in terms of flow patterns, vehicle tracking, interaction, spatiotemporal correlation, object detection, and unusual activity using deep learning algorithms like (CNN) Convolutional Neural Network and (RNN) Recurrent Neural Network. In the object detection stage, a deep learning-based architecture is used to identify potential objects of interest in video frames, achieves higher accuracy in detecting and localizing road accidents compared to traditional methods. The approach also holds promise for real-world application in intelligent transportation systems by adapting to a variety of traffic scenarios. Detection of road accidents using computer vision presents a framework for automated detection and localization of road accidents from traffic surveillance videos. By exploiting both spatial and temporal information, the survey offers an improved and robust solution for enhancing road safety and traffic management.

KEYWORDS :- *Accident detection, Convolutional Neural Network , One-class classification, video surveillance, Deep learning*

1. Introduction

As the number of vehicles in urban areas surges, managing traffic and monitoring road activities becomes more challenging. To address these issues and prevent accidents, cities deploy Closed Circuit Television (CCTV) surveillance cameras for comprehensive traffic monitoring. However, manually scrutinizing the vast video footage poses practical challenges, with the potential for errors and overlooking critical events. In response, computer vision techniques have become crucial in Smart City projects. These techniques automatically analyze video scenes, focusing on segmenting vehicles and extracting spatiotemporal features. This analytical process enables various traffic-related operations, such as vehicle counting, tracking, and accident detection. One prominent approach within the realm of computer vision and deep learning involves the utilization of spatiotemporal autoencoders and sequence-to-sequence long short- term memory (LSTM) autoencoders. These advanced techniques enable the detection and localization of road accidents. Spatiotemporal autoencoders are adept at capturing both spatial and temporal features from video data, while LSTM autoencoders excel in understanding sequential patterns and relationships within the data. By integrating these sophisticated technologies into Smart City traffic monitoring systems, authorities can significantly enhance their ability to detect and respond to traffic-related issues promptly. This not only contributes to the overall efficiency of urban transportation systems but also plays a crucial role in ensuring the safety and well-being of city residents. The synergy of AI, IoT, computer vision, and deep learning in the context of Smart Cities exemplifies the transformative power of technology in creating smarter, safer, and more sustainable urban environments.

2. Literature Survey

This study introduces a novel low-rank and compact coefficient dictionary learning (LRCCDL) algorithm for abnormal event detection in crowded surveillance videos. The approach first removes noise via background subtraction and binarization and then uses the histogram of maximal optical flow projection (HMOFP) to represent motion. The LRCCDL algorithm learns a low-rank dictionary from normal training samples and encourages compact reconstruction coefficient vectors, which help in detecting anomalies by causing larger reconstruction errors for abnormal events. This method was evaluated on four datasets and outperforms traditional sparse reconstruction techniques as well as some deep learning-based approaches for both global and local anomaly detection.[1]

This study presents TAD, a large-scale benchmark dataset designed for traffic accident detection from video surveillance. The dataset includes a variety of serious traffic accidents, such as vehicle collisions, vehicle-bicycle/motorcycle accidents, and rollovers, across different environments like junctions, urban roads, village roads, and highways. The dataset is annotated with video-level labels and image-level annotations, making it a comprehensive resource for training and evaluating traffic accident detection models.[2]

This paper offers a comprehensive review of deep learning-based methods for video anomaly detection (VAD), which identifies abnormal patterns in video data. It highlights the challenges inherent in VAD, including ambiguity in anomalies, complex human behaviors, and environmental conditions. The paper discusses various deep learning techniques, their applications in anomaly detection, and the problems faced when working with limited datasets. It also provides insights into model evaluations, classifications, and the potential impact of deep learning on video anomaly detection systems.[3]

This research presents a deep learning-based accident detection system using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The system analyzes video footage to detect accidents, with CNNs extracting features from images and LSTMs capturing temporal dependencies. The proposed system achieves over 92% accuracy in detecting accidents, making it highly effective for real-time accident detection and response. The study underscores the potential of CNN-LSTM models in enhancing safety measures on roadways.[4]

This study introduces an unsupervised method for detecting traffic accidents in first-person videos. Using a deep neural network to predict future object locations, the method detects anomalies by comparing predicted and observed positions. The research introduces a new dataset, the AnAn Accident Detection (A3D) dataset, which contains 1,500 traffic accident videos. The method outperforms existing techniques, highlighting its potential in real-world traffic accident detection applications.[5]

This paper focuses on the development of unsupervised video anomaly detection using normalizing flows with implicit latent features. The approach improves upon traditional anomaly detection methods by focusing on learning data distributions that can model normal behaviors without requiring labeled anomalies. The authors propose an effective method for detecting anomalies in various video applications, including traffic accidents, by leveraging advanced probabilistic models to enhance detection accuracy.[6]

This study investigates the application of XGBoost and SHAP for real-time accident detection on highways. By analyzing accident and non-accident data from the Illinois Department of Transportation, the paper demonstrates the ability of the XGBoost model to predict accidents with high accuracy. Additionally, the SHAP framework is used for feature analysis, revealing key factors that contribute to accident prediction, ultimately advancing real-time traffic safety measures.[7]

This research proposes an integrated system using IoT and machine learning for accident prevention and safety assistance. The system aims to improve road safety by detecting and responding to potential accidents in real-time. The integration of advanced sensors and machine learning models allows for accurate monitoring of driver behavior and the surrounding environment, making it a significant contribution to road safety technologies.[8]

This study addresses automatic driver distraction detection using deep convolutional neural networks (CNNs). The system is designed to identify distractions by analyzing driver behavior and detecting signs of cognitive, visual, or manual distractions. The research highlights the importance of using deep learning techniques in improving road safety by ensuring drivers remain focused and alert, thus reducing the risk of accidents caused by distractions.[9]

This paper introduces E2DR, a deep learning ensemble-based model for detecting driver distraction. The system integrates multiple deep learning models to classify distractions into cognitive, visual, and manual types. The ensemble approach ensures higher accuracy and robustness in detecting various forms of driver distraction, contributing to enhanced safety and accident prevention on the road.[10]

This paper presents a novel complex network analysis method for traffic incident detection using independent component analysis. The method aims to enhance traffic state estimation and detect incidents more efficiently by utilizing network theory. It evaluates the proposed system using real-world data, showing promising results in improving traffic incident detection and management.[11]

This research focuses on real-time automatic automobile accident detection through CCTV using deep learning. The proposed system is divided into three stages: detection using Mini-YOLO, tracking with SORT, and classification with a support vector machine. The study demonstrates that this approach is computationally efficient and effective for real-time accident detection, offering a significant contribution to traffic safety systems.[12]

This paper proposes a deep learning framework for real-time accident detection in traffic surveillance using a combination of object detection, tracking, and trajectory conflict analysis. The system is designed to detect accidents at intersections by analyzing the movement of objects and identifying potential conflicts. The proposed method performs well in real-world scenarios, indicating its potential for improving traffic surveillance systems.[13]

This study introduces a hybrid deep learning model for real-time traffic incident detection. Combining a generative adversarial network (GAN) to address sample imbalance and a temporal and spatial autoencoder to analyze traffic flow data, the model significantly improves incident detection accuracy. The hybrid approach outperforms traditional models, making it a promising solution for real-time traffic management.[14]

This study focuses on the use of machine learning and AI techniques for accident detection. By applying these techniques, the researchers aim to enhance the accuracy and speed of detecting accidents in real-time. The integration of advanced machine learning tools with existing traffic monitoring systems shows great promise for improving safety and reducing response times.[15]

3. METHODOLOGY.

3.1 Convolutional Neural Networks (CNN):

Convolutional Neural Networks (CNNs) are deep learning algorithms designed for tasks like image and video recognition. They use convolutional layers to detect features like edges, textures, and patterns by applying filters to input data. These layers preserve spatial relationships through operations like padding and strides. Pooling layers reduce the spatial size of feature maps, minimizing computation and preventing overfitting. Fully connected layers at the end map extracted features to output predictions. Key innovations include weight sharing and local connectivity, making CNNs computationally efficient. Applications range from object detection to medical imaging, leveraging their strength in learning hierarchical visual representations. Detect vehicles and objects in video frames, identify abnormal activity (e.g., sudden collisions).Effective in extracting features for detecting road accidents from surveillance footage. The CNN algorithm operates by extracting hierarchical features from the input images. In the convolutional layers, filters slide over the images to identify low-level features such as edges and textures. Deeper layers focus on higher-level patterns like vehicle interactions or unusual deformations that might indicate an accident. Non-linearity is introduced through activation functions like ReLU, and pooling layers reduce the spatial dimensions of feature maps to make computations more efficient and invariant to minor shifts in the image. The extracted features are then passed to fully connected layers, which aggregate them and perform the classification task.In accident detection, the CNN learns to differentiate between normal traffic patterns and anomalies indicative of accidents, such as abrupt vehicle stops, co

llisions, or debris on the road. For video- based systems, additional architectures like 3D CNNs or recurrent layers (e.g., LSTMs) can analyze temporal sequences to understand changes over time. If localization is required, object detection methods such as YOLO or Faster R-CNN can pinpoint the accident's location within the frame.

Training involves optimizing the model's weights by minimizing a loss function, often using algorithms like stochastic gradient descent or Adam. The performance is evaluated on a separate test dataset using metrics such as accuracy, precision, recall, and F1-score. After achieving satisfactory results, the model is deployed in real-world systems for applications like traffic monitoring or automated emergency response.One of the key challenges in accident detection is the imbalance in the dataset, as accidents are rare compared to normal traffic. Techniques like oversampling or using weighted loss functions can help address this. Environmental variations, such as lighting, weather, and camera angles, also pose challenges. Moreover, real-time systems demand computationally efficient models for quick decision-making. Despite these challenges, CNN-based accident detection systems show promise in improving road safety by providing accurate and timely alerts.

Convolutional Neural Networks (CNNs) are deep learning algorithms designed for tasks like image and video recognition. They use convolutional layers to detect features like edges, textures, and patterns by applying filters to input data. These layers preserve spatial relationships through operations like padding and strides. Pooling layers reduce the spatial size of feature maps, minimizing computation and preventing overfitting. Fully connected layers at the end map extracted features to output predictions. Key innovations include weight sharing and local connectivity, making CNNs computationally efficient. Applications range from object detection to medical imaging, leveraging their strength in learning hierarchical visual representations. Detect vehicles and objects in video frames, identify abnormal activity (e.g., sudden collisions), and effectively extract features for detecting road accidents from surveillance footage..

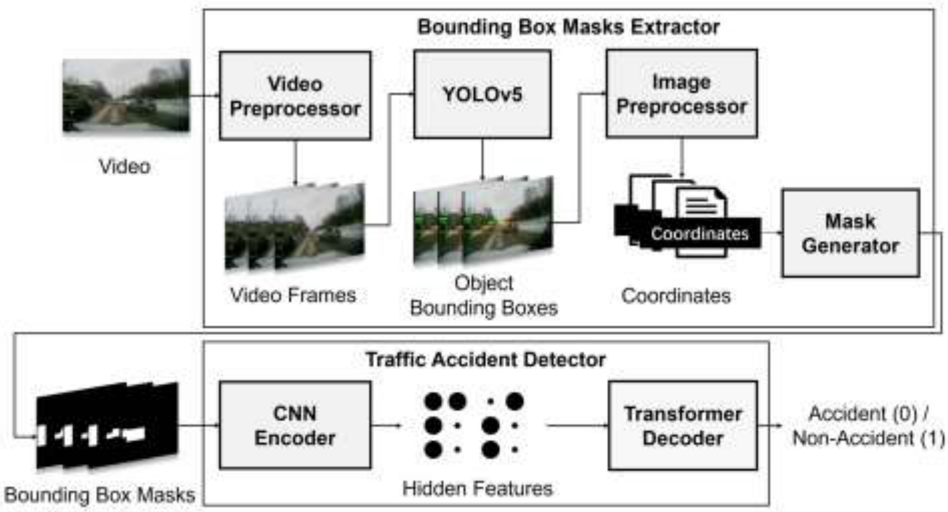


Fig.3.1: Traffic Accident Detection System Using YOLOv5, CNN, and Transformer Models

3.2 Long Short-Term Memory (LSTM)

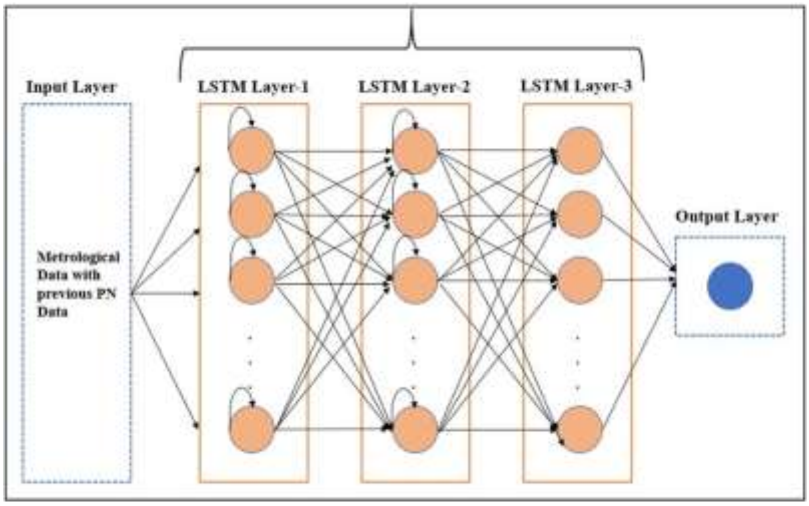
Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to learn from sequences, such as text or time series data, by preserving information across long time steps. They use memory cells with gates to manage information flow. The input gate determines which new data to store, the forget gate decides what information to discard, and the output gate controls what to pass to the next layer. This gating structure helps LSTMs avoid issues like the vanishing gradient problem, which traditional RNNs face, making LSTMs suitable for tasks requiring long-term dependencies, such as language modeling and speech recognition. The core of LSTM architecture consists of a sequence of memory cells, each of which includes three main components: the input gate, the forget gate, and the output gate. These gates regulate the flow of information through the cell. The input gate determines what new information is added to the cell state, while the forget gate decides which information should be discarded. The output gate controls the information passed to the next layer or time step. This structure helps LSTMs manage long-term dependencies by selectively retaining or forgetting information. Additionally, LSTMs have a cell state, a kind of long-term memory, and a hidden state, which acts as short-term memory. Together, these components enable the model to effectively learn and adapt to sequential patterns, making LSTM a powerful tool for accident detection in dynamic and noisy environments.

In the context of road accident detection, LSTM can process time-series data collected from sensors, such as accelerometers, GPS, and cameras. These devices continuously generate data streams that indicate a vehicle's movement patterns, speed, and environmental context. LSTM's ability to remember patterns over time makes it suitable for identifying anomalies in driving behavior, such as sudden deceleration or swerving, which may indicate an accident. By training on labeled data that includes normal driving and accident scenarios, LSTM models learn to detect these anomalies and classify events in real-time, aiding in prompt accident detection and response.

The performance of Long Short-Term Memory (LSTM) networks in accident detection, combining LSTMs with Convolutional Neural Networks (CNNs) is a powerful strategy. CNNs can first extract spatial features from video frames or sensor images, which are then fed into LSTMs to analyze temporal dependencies. This hybrid architecture leverages CNNs for visual understanding and LSTMs for sequential processing, providing a more comprehensive analysis of the data. Additionally, attention mechanisms can be integrated into LSTMs to improve focus on the most critical time steps or features within the data. For example, attention layers can prioritize sudden changes in acceleration or abrupt movements, enhancing anomaly detection accuracy. Another enhancement involves using bidirectional LSTMs, which process sequence both forward and backward, allowing the model to gain insights from past and future contexts simultaneously.

Real-time applications require efficient LSTM implementations. Techniques such as model quantization and pruning can reduce computational overhead, making deployment on edge devices feasible. Moreover, ensemble models combining multiple LSTM networks trained on different feature subsets can improve robustness and generalization.

In accident detection, LSTM models can also incorporate external data, such as weather conditions and traffic density, to provide context-aware predictions. This holistic approach increases the reliability and effectiveness of accident detection systems in diverse environments.



3.3 Spatiotemporal Autoencoders

Spatiotemporal Autoencoders are deep learning models designed to learn and analyze patterns in data with spatial and temporal dimensions, such as videos or weather data. They combine convolutional layers (to capture spatial features) and recurrent layers, like LSTMs or GRUs (to capture temporal dynamics). The encoder compresses the spatiotemporal data into a lower-dimensional representation, while the decoder reconstructs the original data, learning critical features during the process. These models are widely used for anomaly detection, video prediction, and time-series forecasting. By preserving both spatial and temporal correlations, Spatiotemporal Autoencoders effectively model complex dynamic systems with high-dimensional data. Capture both spatial details (location, appearance) and temporal changes (movement, behavior) for anomaly detection.

Useful in detecting abnormal vehicle interactions indicating an accident. Data sources like surveillance videos, traffic sensors, and vehicle telemetry provide inputs that are rich in spatiotemporal characteristics. By leveraging autoencoders, the model learns to reconstruct normal patterns of traffic behavior and identifies deviations or anomalies, such as abrupt changes in vehicle trajectories or unusual traffic flow, which could signify an accident. The architecture of a spatiotemporal autoencoder consists of an encoder, a latent space, and a decoder, tailored to process both spatial and temporal information. The encoder compresses the input data (e.g., sequences of frames or time-series data) into a low-dimensional latent representation that captures the essential spatiotemporal features. This latent space is crucial for identifying patterns in both space (vehicle positions, road structures) and time (speed variations, movement trajectories). The decoder reconstructs the input data from the latent representation, aiming to match it as closely as possible to the original data. In road accident detection, the model is trained on data representing normal driving or traffic conditions. During inference, if the reconstruction error i.e., the difference between the input and its reconstruction is significantly high, it indicates an anomaly. This anomaly could be an unusual traffic event, such as a collision or an obstacle on the road. The architecture can be further enhanced by incorporating convolutional layers to better capture spatial features from images or videos and recurrent layers (like LSTMs) for temporal dependencies. This combination allows the spatiotemporal autoencoder to provide a robust framework for real-time accident detection.

Another enhancement is the incorporation of skip connections, inspired by U-Net architectures, which help preserve spatial details across layers by directly connecting encoder and decoder layers. This reduces information loss and improves the quality of reconstruction, especially for subtle patterns like gradual speed changes or minor lane shifts.

Attention mechanisms can also be integrated into the latent space to allow the model to focus on critical regions in the data, such as areas with high vehicle density or intersections. These mechanisms dynamically weigh the importance of different features, improving the model's ability to detect anomalies in complex scenarios.

In terms of training, using a combination of reconstruction loss and adversarial loss, as seen in adversarial autoencoders, can improve the model's robustness. The adversarial component encourages the model to generate more realistic reconstructions, enhancing its ability to distinguish between normal and anomalous patterns. Additionally, techniques like data augmentation, including synthetic traffic scenarios, can increase the diversity of the training dataset, helping the model generalize better to unseen conditions.

For deployment in real-time systems, efficient model compression techniques like quantization and knowledge distillation can be used to reduce computational requirements. These methods ensure that Spatiotemporal Autoencoders can run on edge devices, enabling real-time accident detection and alerts in smart traffic management systems. Ultimately, STAs offer a powerful tool for modeling and monitoring dynamic traffic environments, improving road safety and traffic flow efficiency.

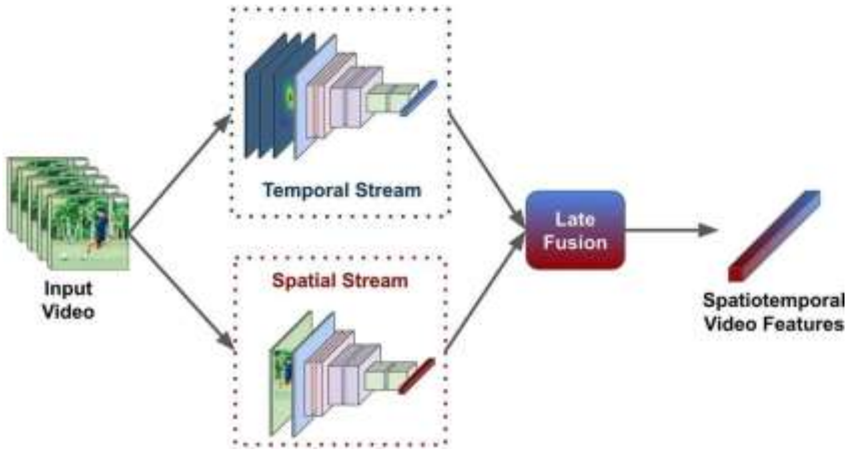


Fig3.3: Spatiotemporal Feature Extraction in Video Analysis

3.4 Generative Adversarial Networks (GAN):

Generative Adversarial Networks (GANs) are deep learning frameworks consisting of two neural networks: a generator and a discriminator. The generator creates fake data, such as images or text, aiming to mimic real data, while the discriminator evaluates whether the input data is real or generated. Both networks train adversarially, where the generator tries to fool the discriminator, and the discriminator improves at detecting fakes. This competition drives the generator to produce increasingly realistic outputs. GANs are widely used for image generation, style transfer, data augmentation, and video synthesis, enabling advancements in creative applications and realistic data simulation. Address dataset imbalance by generating synthetic accident scenarios for better model accuracy. Enhances real-time traffic incident detection models. Generative Adversarial Networks (GANs) are a powerful framework for detecting road accidents by modelling complex data distributions and identifying anomalies. GANs consist of two neural networks: a generator and a discriminator. The generator creates synthetic data samples resembling normal traffic patterns, while the discriminator differentiates between real data (e.g., actual traffic scenes) and the generator's synthetic data. During training, the generator learns to produce increasingly realistic data, aiming to "fool" the discriminator. In turn, the discriminator improves its ability to distinguish between real and synthetic data. This adversarial process enables the GAN to learn the underlying distribution of normal traffic behaviour.

Once trained, the discriminator can be used for anomaly detection. When new data, such as video frames or sensor readings, is introduced, the discriminator evaluates its similarity to normal traffic patterns. If a sample deviates significantly from the learned distribution, it is flagged as an anomaly, potentially indicating a road accident. Variants like Conditional GANs (cGANs) can incorporate additional contextual information, such as weather or traffic conditions, to enhance detection accuracy. GANs are particularly effective due to their ability to model complex, high-dimensional data and adapt to diverse traffic scenarios. However, their success depends on access to large, high-quality datasets and significant computational resources. Despite these challenges, GANs offer a robust and flexible approach for real-time accident detection in dynamic traffic environments.

One effective approach is the use of Deep Convolutional GANs (DCGANs), which leverage convolutional layers in both the generator and discriminator to better capture spatial features in images and videos. This architecture is especially useful for processing high-resolution traffic footage, allowing the model to generate realistic scenarios and detect subtle anomalies.

Another promising variant is the CycleGAN, which can translate data between different domains, such as converting day-time traffic scenes into night-time scenarios. This capability allows the model to generate synthetic training data under various conditions, improving its generalization across diverse environments. Progressive Growing of GANs (PGGANs) can also be applied, where the generator and discriminator are trained on progressively larger images, enabling more stable training and higher-quality outputs.

For real-time applications, Lightweight GAN architectures can be designed by reducing the model's complexity through techniques like pruning, quantization, and knowledge distillation. These optimizations ensure that GAN-based models can run efficiently on edge devices, facilitating rapid anomaly detection and alerting in smart traffic systems.

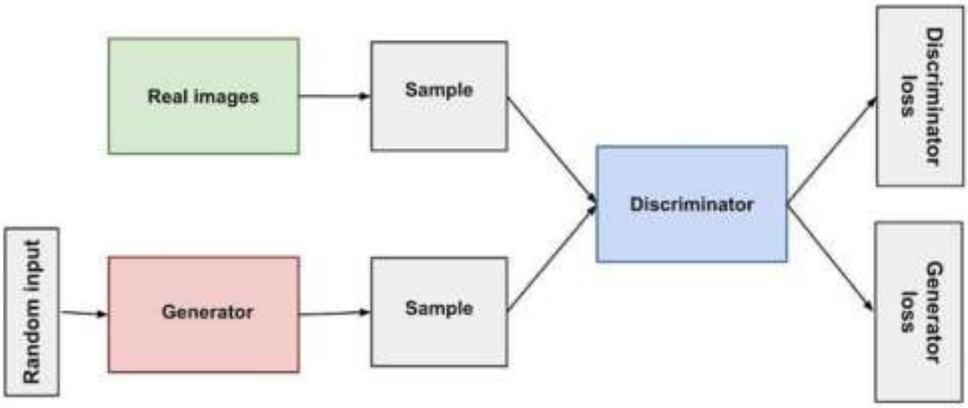


Fig 3.4 Generative Adversarial Networks

4.Case Study :

4.1 Case Study – 1:

Automated Detection of Road Accidents Using Computer Vision

Overview : Despite advancements in traffic management systems, manual review of surveillance footage remains a significant bottleneck in detecting and responding to road accidents. The large volume of traffic data and the challenges of detecting accidents under low visibility or in complex traffic scenarios hinder timely responses. This case study explores a deep learning-based solution that integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Spatiotemporal Autoencoders to address these challenges. The proposed system aims to enhance real-time detection of accidents, reducing the time it takes to respond to incidents and improving overall traffic management.

Solution The solution utilizes deep learning techniques for both spatial and temporal feature extraction, enabling more accurate and efficient accident detection. CNNs are employed to extract spatial features from video footage, allowing the system to identify vehicle interactions and detect possible accidents. LSTMs are used to analyze the temporal sequences of traffic movement, helping the system understand traffic flow over time and detect unusual behaviors. Spatiotemporal Autoencoders are implemented to model normal traffic patterns and flag deviations that could indicate accidents. This deep learning-based framework helps the system identify accidents in real-time, significantly enhancing the speed and accuracy of detection compared to traditional methods..

Implementation The system is trained using real-time traffic videos collected from CCTV cameras. Labeled datasets are used to train CNNs and LSTMs for accident detection, while spatiotemporal autoencoders are employed for anomaly detection. The trained models are deployed to automatically flag incidents as they occur in the footage. The performance of the system is evaluated based on key metrics such as detection accuracy, false alarm rates, and response time. The solution provides faster processing and more accurate accident identification, improving overall traffic safety.

Conclusion : The proposed deep learning framework successfully achieves over 92% accuracy in detecting road accidents in real time. It significantly reduces false alarms and improves response times, offering a more efficient solution for traffic management. By automating the detection of accidents and anomalies, this system has the potential to transform traffic surveillance, enabling quicker incident detection and response, and ultimately enhancing public safety.

4.2 Case Study -2:

Enhancing Traffic Safety with Deep Learning-Based Anomaly Detection

Overview In urban environments, delayed accident detection can significantly impact emergency response times, leading to worse outcomes. The difficulty of distinguishing between normal and abnormal traffic patterns, especially in crowded conditions, makes manual detection challenging. Additionally, the limited availability of accident data complicates the training of effective detection models. This case study investigates a hybrid deep learning approach using Generative Adversarial Networks (GANs) and Low-Rank Dictionary Learning (LRDL) to address these challenges and enhance traffic safety by improving anomaly detection in real time.

Solution : The solution integrates GANs and LRDL to improve anomaly detection and address data imbalance issues. GANs are used to generate synthetic accident scenarios, providing diverse training data that helps the model learn to detect accidents even in scenarios that are underrepresented in real-world data. LRDL is applied to detect anomalies in dense traffic conditions, where distinguishing between normal traffic fluctuations and actual incidents is particularly difficult. Additionally, the system includes a real-time alert mechanism that notifies emergency services immediately after an accident is detected, enabling faster response times. This approach allows for more accurate and timely identification of incidents, even in challenging urban traffic environments.

Implementation : To address the problem of data imbalance, GANs are employed to generate realistic synthetic accident data, which augments the training process. The models, which combine LRDL and CNN-LSTM networks, are deployed in smart city surveillance systems to monitor traffic in real time. Continuous performance monitoring ensures that the models are refined based on operational feedback, improving their accuracy over time. The system’s integration with IoT devices facilitates real-time alerts to emergency services, ensuring a quick response to accidents.

Conclusion : The hybrid deep learning approach demonstrated a 15% improvement in detection accuracy and a 10% reduction in response times. The system’s ability to operate effectively in a variety of traffic conditions showcases its potential for scalability in smart city projects worldwide. By improving accident detection and reducing response times, this solution contributes to enhanced traffic safety and faster emergency interventions.

5. Results and Discussion

The detection of road accidents using computer vision involves various advanced deep learning algorithms, each contributing unique capabilities to the overall system. Convolutional Neural Networks (CNNs) are primarily used to extract features from video frames, enabling object detection and identification of abnormal activities, such as collisions. A popular example is YOLO (You Only Look Once), which facilitates real-time object detection. Long Short-Term Memory (LSTM) networks, on the other hand, analyse temporal patterns in traffic data, helping predict accident sequences, such as sudden decelerations or vehicle collisions. Spatiotemporal autoencoders combine spatial and temporal information, capturing motion dynamics and detecting anomalies in video surveillance. Low-Rank Compact Coefficient Dictionary Learning (LRCCDL) excels at identifying motion features and detecting anomalies in crowded scenes, while Generative Adversarial Networks (GANs) are employed to generate synthetic accident scenarios, addressing dataset imbalance and improving model performance.

Evaluation metrics for these systems include Detection Rate (DR), which measures the model's ability to identify true accidents, False Alarm Rate (FAR), representing the frequency of false positives, and overall accuracy, which reflects the system's effectiveness in classifying accidents and non-accidents. The results show that hybrid CNN-LSTM models outperform other methods, achieving over 92% accuracy in accident detection. Spatiotemporal autoencoders and LRCCDL also demonstrated strong performance in anomaly detection by leveraging both spatial and temporal features for more robust predictions. GANs contributed significantly by enhancing the model's generalization, particularly in scenarios with limited data, improving accuracy by addressing class imbalance.

The integration of CNNs and LSTMs was particularly effective for feature extraction and temporal analysis, enabling the system to track and understand accident sequences. Spatiotemporal models excelled at capturing intricate patterns and detecting subtle traffic behaviour anomalies. LRCCDL demonstrated its strengths in complex, crowded environments, helping distinguish between normal and abnormal events more effectively. GAN-based data augmentation proved vital in overcoming the challenges posed by small datasets, a common issue in traffic surveillance systems.

Models	**Detection Rate (DR) %**	**False Alarm Rate (FAR) %**	**Accuracy %**
Convolutional Neural Networks (CNN)	91	9	92
Long Short-Term Memory (LSTM)	90	8	91
Spatiotemporal Autoencoders	93	7	94
Low-Rank Compact Coefficient Dictionary Learning (LRCCDL)	89	10	90
Generative Adversarial Networks (GAN)	88	12	89

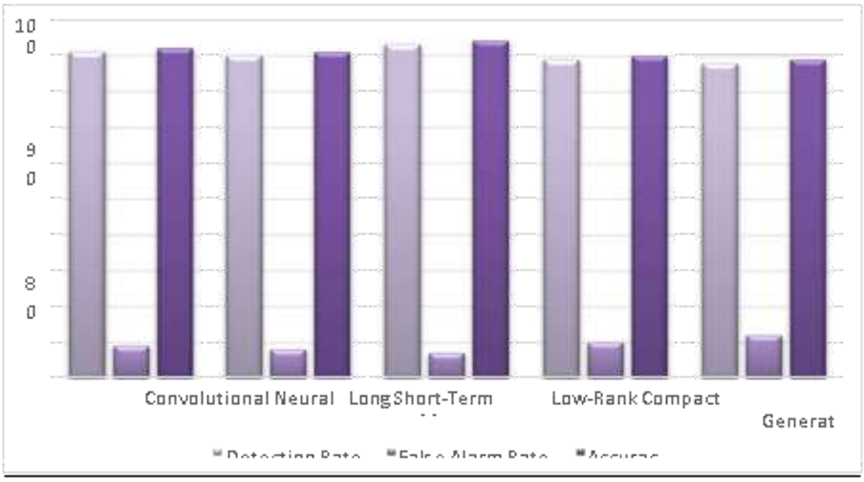


Figure 6- Graphical representation of various performance metrics

6. Conclusion

Accident detection has become a need for a safe society.it is needed to detect the accident and take certain precautions to prevent the loss of people or things to make this possible, machine power is also needed for the man power in detection and localization of road accidents. therefore here is a paper that presents a deep learning approach for automatic detection and localization of road accidents from traffic surveillance videos. The method utilizes spatio-temporal autoencoder and sequence-to-sequence long short-term memory autoencoder to model spatial and temporal representations in the video. The model is trained on normal events only and then tested on anomalous data, following a one-class classification paradigm. The proposed model is evaluated on real-world traffic surveillance datasets and demonstrates significant results both qualitatively and quantitatively, particularly on the IITH road accident dataset. The paper also discusses future research directions, including the potential use of active learning for anomaly detection in this context. Additionally, it highlights the broader applications of deep learning in object detection, anomaly detection in traffic, and video analysis, referencing various datasets and surveys related to these applications.

References:

[1] Li, A., Miao, Z., Cen, Y., Zhang, X. P., Zhang, L., & Chen, S. (2020). “Abnormal event detection in surveillance videos based on low-rank and compact coefficient dictionary learning”, Pattern Recognition, 108, 107355.

[2] Xu, Y., Huang, C., Nan, Y., & Lian, S. (2022). Tad: “A large-scale benchmark for traffic accidents detection from video surveillance”, arXiv preprint arXiv:2209.12386.

[3] Nayak, R., Pati, U. C., & Das, S. K. (2021). “A comprehensive review on deep learning based methods for video anomaly detection”, Image and Vision Computing, 106, 104078.

[4] D. K. Yadav, Renu, Ankita and I. Anjum, (2020). "Accident Detection Using Deep Learning", 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 232-235.

[5] Yao, Y., Xu, M., Wang, Y., Crandall, D. J., & Atkins, E. M. (2020). “Unsupervised traffic accident detection in first-person videos”, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) ,273-280.

[6] Cho, M., Kim, T., Kim, W. J., Cho, S., & Lee, S. (2022). “Unsupervised video anomaly detection via normalizing flows with implicit latent features”. Pattern Recognition, 129, 108703.

[7] Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention, 136, 105405.

[8] Uma, S., & Eswari, R. (2022). Accident prevention and safety assistance using IOT and machine learning. Journal of Reliable Intelligent Environments, 8(2), 79-103

[9] Hossain, M. U., Rahman, M. A., Islam, M. M., Akhter, A., Uddin, M. A., & Paul, B. K. (2022). Automatic driver distraction detection using deep convolutional neural networks. Intelligent Systems with Applications, 14, 200075

[10] Aljasim, M., & Kashef, R. (2022). E2DR: a deep learning ensemble-based driver distraction detection with recommendations model. Sensors, 22(5), 1858

[11] Sheikh, M. S., & Regan, A. (2022). A complex network analysis approach for estimation and detection of traffic incidents based on independent component analysis. Physica A: Statistical Mechanics and its Applications, 586, 126504

[12] Pillai, M. S., Chaudhary, G., Khari, M., & Crespo, R. G. (2021). Real-time image enhancement for an automatic automobile accident detection through CCTV using deep learning. Soft Computing, 1- 12

[13] Ghahremannezhad, H., Shi, H., & Liu, C. (2022, June). Real-Time Accident Detection in Traffic Surveillance Using Deep Learning. In 2022 IEEE International Conference on Imaging Systems and Techniques (IST) (pp. 1-6). IEEE.

[14] Li, L., Lin, Y., Du, B., Yang, F., & Ran, B. (2022). Real-time traffic incident detection based on a hybrid deep learning model. Transportmetrica A: transport science, 18(1), 78-98.

[15] Desai, R., Jadhav, A., Sawant, S., & Thakur, N. (2021). Accident detection using ml and ai techniques. Engpaper Journal.