ACCIDENT DETECTION WITH CCTV USING CONVOLUTIONAL NEURAL NETWORK

[1]K Murugan, [2]Charan Gopi Yerramsetty, [3]Naresh Kumar Tadimari, [4]Sai Bhargav Tadimari, [5]Pavan Kumar Sanku

[1]Assistant Professor, Department of Computer Science and Engineering,

Hindusthan Institute of Technology, Coimbatore

[1]*murugan.k@hit.edu.in*

2,3,4,5 UG Students, Department of Computer Science and Engineering,

Hindusthan Institute of Technology, Coimbatore

[2]charangopi19@gmail.com,[3]tadimari_info@gmail.com,[4]saibhargavcse@gmail.com,[5]sankubecse@gmail.com

Abstract Accident detection is essential for ensuring timely emergency response, reducing fatalities, and minimizing property damage. This project presents an automated accident detection system that utilizes Closed-Circuit Television (CCTV) footage combined with Convolutional Neural Networks (CNNs) to identify accidents on roads in real time. CCTV cameras installed at key traffic locations continuously capture video footage, which is analyzed frame by frame using a CNN model trained on a comprehensive dataset of annotated accident and non-accident scenarios. The model learns to recognize patterns indicative of accidents, such as sudden changes in vehicle motion, anomalies in traffic flow, and the presence of emergency vehicles. Once trained, the system can accurately detect potential accidents from live footage and immediately trigger alerts to emergency services, providing vital details such as the location and nature of the incident. In addition to notifying authorities, the system can also be configured to send warnings to nearby vehicles and pedestrians, enabling them to take precautionary actions or reroute to avoid the affected area. This approach eliminates the dependency on human operators, who may miss critical events due to fatigue or limited attention, and ensures a faster, more reliable response. By integrating artificial intelligence with existing CCTV infrastructure, the system offers a scalable and cost-effective solution for improving road safety and traffic monitoring. Future enhancements may include incorporating additional sensor inputs—such as vehicle speed, traffic density, and weather conditions—to improve detection accuracy and enable predictive accident prevention. Overall, the proposed system demonstrates a practical and impactful application of deep learning in intelligent transportation systems.

Keywords: Convolutional Neural Network; Accident Detection; Deep Learning; Video Classification; Recurrent Neural Network.

1. INTRODUCTION

Road traffic accidents are one of the leading causes of death and serious injury worldwide. According to the World Health Organization (WHO), approximately 1.3 million people die each year due to road traffic crashes, and millions more suffer non-fatal injuries, often with life-altering consequences. These incidents not only have a devastating human cost but also pose a significant economic burden due to lost productivity, medical expenses, and property damage. With the increasing number of vehicles on roads and growing urbanization, the frequency and severity of road accidents have been escalating, necessitating the development of effective systems to detect and respond to such events swiftly.

Timely detection of road accidents is critical for initiating rapid emergency responses, which can significantly reduce casualties, mitigate traffic congestion, and prevent secondary collisions. Traditional accident detection methods largely depend on manual monitoring of CCTV footage by human operators or on reports from passersby and drivers. These methods, while functional, are slow, inefficient, and prone to error. Human operators

may miss crucial events due to fatigue, distraction, or the overwhelming volume of video feeds they need to monitor. In many cases, accidents may go unnoticed for several minutes, delaying the arrival of emergency services and increasing the risk of additional incidents.

With advancements in artificial intelligence and computer vision, there is a growing interest in automating the process of accident detection using machine learning techniques. Among these, Convolutional Neural Networks (CNNs) have emerged as a powerful tool for analyzing visual data. CNNs are a class of deep learning algorithms that excel at extracting features from images and video frames, enabling them to identify complex patterns and anomalies. They have been successfully applied in various domains, such as facial recognition, medical imaging, and autonomous driving. In the context of traffic monitoring, CNNs can be trained to distinguish between normal and abnormal traffic behavior, making them well-suited for real-time accident detection.

This project proposes a system that leverages CCTV footage and CNNs to automatically detect road accidents. The system utilizes existing CCTV infrastructure commonly deployed at intersections, highways, and urban road networks. Real-time video streams from these cameras are fed into a CNN model, which analyzes the frames to identify features indicative of accidents, such as abrupt deceleration, unusual vehicle trajectories, or sudden changes in traffic flow. The model is trained on a labeled dataset comprising accident and non-accident video clips, allowing it to learn visual cues associated with collision events. Once an accident is detected, the system triggers an alert that is sent to relevant authorities, such as emergency responders or traffic control centers, enabling a prompt and informed response.

The motivation behind this project is to create a proactive, scalable, and reliable accident detection system that reduces reliance on human intervention while enhancing road safety and emergency response efficiency. The use of CNNs in this context offers several advantages. First, CNNs can learn high-level features from raw video data without the need for handcrafted feature extraction, making the system more adaptable and robust. Second, the ability to process and analyze video in real time ensures that accidents are detected as soon as they occur, minimizing delays in reporting. Third, since the system uses already-installed CCTV cameras, it is cost-effective and can be integrated into existing smart city and intelligent transportation infrastructure.

Existing solutions for accident detection include sensor-based methods, such as using accelerometers, GPS, and onboard diagnostics in vehicles, as well as vision-based approaches involving traditional image processing techniques. Sensor-based methods, while accurate, require installation in every vehicle and cannot detect accidents involving unmonitored vehicles or pedestrians. Traditional image processing methods, such as background subtraction and motion estimation, often struggle with occlusions, lighting changes, and complex traffic scenarios. In contrast, CNN-based methods offer improved performance in these conditions due to their ability to learn from large and diverse datasets.

Despite the promise of CNNs in accident detection, there are several challenges that must be addressed. One major challenge is the availability and quality of training data. Real accident footage is relatively rare, and collecting a comprehensive dataset that captures various types of accidents under different environmental conditions is difficult. Data augmentation and simulation techniques can help expand the dataset, but they may not fully replicate real-world complexities. Another challenge is ensuring the system's robustness to different camera angles, lighting conditions, weather variations, and traffic densities. To address this, the CNN model must be trained and validated on a diverse set of video samples that reflect real-world variability.

Moreover, achieving real-time performance is critical for the practical deployment of such a system. The model must be efficient enough to process high-resolution video streams without significant latency. This may require optimization techniques such as model pruning, quantization, or the use of specialized hardware accelerators like GPUs or TPUs. Additionally, false positives and false negatives must be minimized to avoid unnecessary alarms or missed detections, which could undermine trust in the system. Balancing accuracy, speed, and reliability is a key objective of this research. The broader impact of this project lies in its contribution to intelligent transportation systems (ITS) and smart city initiatives. As urban centers strive to become more connected and data-driven, the integration of AI-powered accident detection systems into traffic management

frameworks can significantly improve safety, reduce congestion, and enhance the overall efficiency of road networks. Furthermore, the data collected by such systems can be used for traffic analysis, urban planning, and policy-making, leading to more informed decisions and safer infrastructure design.

2. Literature Survey

The development of accident detection systems using CCTV footage and Convolutional Neural Networks (CNNs) has gained significant attention in recent years. This section reviews several notable studies that have contributed to this field, focusing on their methodologies, datasets, results, and limitations, and highlighting how they relate to the proposed project.

1. Real-Time Accident Detection Using CCTV and Nearest Medical Amenities Suggestion

Pai et al. (2023) introduced a real-time accident detection system leveraging CCTV footage combined with an alert mechanism that suggests nearby medical facilities. The study primarily focuses on using computer vision techniques to identify sudden anomalies in traffic flow that could indicate an accident. Their model utilizes frame-based analysis and extracts motion patterns to detect abnormal events. While the system demonstrates promising results in timely accident detection, it relies heavily on handcrafted features rather than deep learning architectures, which may limit its adaptability to complex scenarios. Moreover, the medical amenities suggestion module highlights a crucial post-detection application, emphasizing the importance of integrated emergency response systems. This work underscores the benefits of coupling accident detection with immediate actionable insights, a concept that aligns well with the proposed system's alerting functionality.

2. Deep Learning Methods for Road Accident Detection Using Traffic Images

Sherimon et al. (2024) developed a deep learning-based approach for accident detection using CNNs trained on annotated traffic images. Their system performs classification to distinguish between accident and non-accident scenes, leveraging convolutional layers to automatically extract relevant features such as vehicle deformation, smoke, or sudden vehicle stops. The authors experimented with various CNN architectures, including VGGNet and ResNet, to optimize accuracy. Their findings indicate that CNN-based methods outperform traditional image processing techniques, particularly in complex traffic environments with occlusions and lighting variability. However, their model is mainly tested on still images rather than continuous video streams, which may affect real-time applicability. The use of pre-trained models and transfer learning techniques presents an efficient way to train the system on relatively small datasets, a technique that could be beneficial for the proposed system given the scarcity of labeled accident videos.

3. CNN-Based Accident Detection System

Aswini et al. (2022) proposed a CNN-based accident detection framework that processes real-time CCTV video footage to identify accidents. Their system focuses on detecting sudden vehicle deceleration and abnormal traffic behavior using spatiotemporal features extracted by a 3D CNN architecture. The model's capability to analyze sequences of frames rather than individual images allows it to capture dynamic changes that signify accidents more effectively. The authors reported high accuracy and low false-positive rates on their test dataset, which includes varied traffic conditions and accident types. This study demonstrates the importance of temporal information in accident detection and justifies the selection of CNN architectures capable of learning from both spatial and temporal data. This approach closely resembles the proposed method and serves as a benchmark for evaluating the performance of other CNN-based models.

4. Deep Learning for Accident Detection Under Challenging CCTV Conditions

Lee and Shin (2019) tackled the challenge of detecting accidents in tunnels, where CCTV monitoring conditions are poor due to low lighting and occlusions. Their approach combines CNNs with data enhancement

techniques to improve robustness under adverse visual conditions. The model is trained on a dataset including both accident and non-accident footage captured under different lighting and weather conditions. Their results emphasize that pre-processing and data augmentation significantly improve CNN performance in challenging environments. This study addresses a common real-world problem that the proposed system may face, such as varying weather and lighting conditions, suggesting that data diversity and augmentation are crucial for improving detection accuracy.

3. Proposed System

The proposed system aims to automatically detect road accidents using real-time CCTV footage by leveraging Convolutional Neural Networks (CNNs). This methodology outlines the overall framework, detailing each component from data acquisition to alert generation. The key focus is on extracting meaningful features from video streams, accurately identifying accidents, and enabling rapid emergency response through automated notifications.

1. System Overview

The architecture consists of four main modules: Data Acquisition, Preprocessing, Accident Detection via CNN, and Alert Generation. CCTV cameras positioned at strategic traffic locations continuously capture video streams. These streams undergo preprocessing to extract frames, enhance image quality, and prepare inputs suitable for CNN analysis. The CNN model then processes the frames or frame sequences to detect accident events by classifying traffic scenarios as either normal or accident-related. Upon detection, an alert is triggered and transmitted to relevant emergency response units and, optionally, to nearby vehicles or pedestrians.
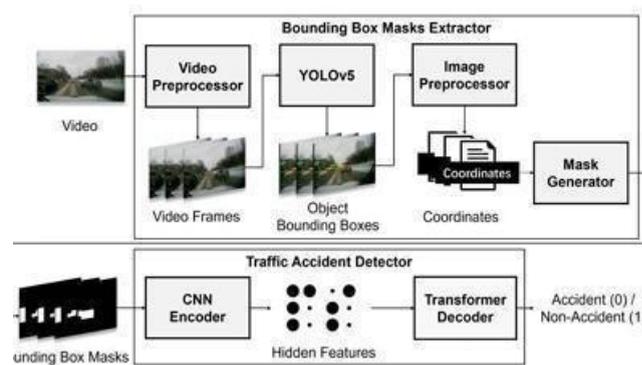
2. Data Acquisition

Accurate and robust accident detection requires high-quality training data that captures a variety of accident scenarios under diverse conditions. Data acquisition involves gathering video footage from multiple CCTV cameras placed at busy intersections, highways, and urban roads. Publicly available datasets, such as the Traffic Accident Dataset (TAD), and real-world CCTV footage from traffic monitoring agencies are collected. The dataset includes videos depicting different types of accidents, such as collisions, sudden stops, and vehicle rollovers, alongside normal traffic scenes for comparison.

Each video is annotated manually or semi-automatically, labeling segments containing accident events. This annotated dataset forms the basis for supervised learning, enabling the CNN to learn discriminative features that differentiate accident from non-accident situations.

3. Preprocessing

Preprocessing is critical to improve model performance by standardizing input data and enhancing feature extraction. The steps include:

- **Frame Extraction:** Continuous video streams are segmented into frames at a fixed rate (e.g., 10-15 frames per second). This rate balances temporal resolution with computational load.
- **Resizing and Normalization:** Each frame is resized to a fixed dimension (e.g., 224x224 pixels) compatible with the CNN input layer. Pixel values are normalized to a range [0, 1] or standardized using mean and standard deviation values from the training set.
- **Noise Reduction and Enhancement:** Techniques such as Gaussian smoothing or histogram equalization may be applied to reduce noise and improve contrast, especially in low-light or adverse weather conditions.
- **Data Augmentation:** To increase dataset variability and improve generalization, augmentation methods such as rotation, flipping, cropping, brightness adjustment, and synthetic occlusion are applied during training.

4. CNN Model Design

The core of the accident detection system is the CNN model designed to extract spatial and temporal features from video frames.

- **Architecture Selection:** A 3D CNN architecture is preferred as it can process sequences of frames (i.e., video clips) to capture motion and temporal dynamics, which are crucial for detecting accidents characterized by sudden changes in vehicle movement. Popular architectures such as C3D or ResNet-3D can be fine-tuned or customized based on dataset specifics.
- **Input Representation:** The model receives clips consisting of a fixed number of consecutive frames (e.g., 16 frames per clip) to analyze temporal progression. Each input tensor has dimensions corresponding to the number of frames, height, width, and color channels.
- **Convolutional Layers:** These layers learn hierarchical spatial features such as vehicle shapes, road patterns, and contextual objects. The 3D convolutions additionally capture temporal variations in pixel intensities, representing movement.
- **Pooling Layers:** Used to downsample feature maps, reducing spatial dimensions while retaining critical information, enhancing computational efficiency.
- **Fully Connected Layers:** These layers aggregate learned features to perform classification into accident and non-accident categories.
- **Activation Functions:** ReLU activations introduce non-linearity, enabling the model to learn complex feature representations.
- **Regularization:** Techniques like dropout and batch normalization prevent overfitting and improve convergence.

5. Model Training

Training the CNN involves feeding annotated video clips and iteratively optimizing the model parameters.

- **Optimization:** The Adam optimizer is commonly chosen for its adaptive learning rate and efficient convergence.
- **Batch Size and Epochs:** Mini-batch training with sizes like 16 or 32 balances memory usage and gradient stability. The model is trained over multiple epochs until performance metrics stabilize.
- **Validation:** A portion of the dataset is reserved for validation to monitor overfitting and tune hyperparameters.

- **Transfer Learning:** Pretrained weights from large-scale video datasets (e.g., Kinetics) can be fine-tuned on the accident dataset to accelerate learning and improve accuracy.
- **Early Stopping:** Training halts when validation loss ceases to improve, preventing overfitting.

6. Real-Time Detection Pipeline

After training, the model is deployed in a real-time environment.

- **Continuous Frame Capture:** Video frames from CCTV cameras are streamed to the detection system.
- **Sliding Window:** Clips of fixed frame length (matching training input) are extracted continuously with overlap to ensure no accident goes undetected.
- **Inference:** Each clip is passed through the CNN, which outputs a probability score indicating the presence of an accident.
- **Thresholding:** A threshold determines whether the clip is classified as an accident. Threshold tuning balances sensitivity and false alarms.
- **Post-Processing:** To reduce noise and false positives, temporal smoothing methods such as majority voting over consecutive clips or temporal filtering can be applied.

7. Alert Generation and Notification

Upon accident detection, the system generates an alert with critical information:

- **Location:** Derived from the specific CCTV camera's geolocation or traffic segment.
- **Time Stamp:** Records the exact time of the event.
- **Nature of Incident:** Based on classification confidence and additional contextual data (e.g., severity inferred from suddenness of motion change).

Alerts are automatically transmitted to emergency services, traffic management centers, and optionally to nearby vehicles or pedestrian notification systems via connected infrastructure or mobile networks. This rapid communication facilitates quicker medical and police response and helps prevent secondary accidents by warning other road users.

RESULTS AND DISCUSSION

This section presents the experimental results obtained from implementing the proposed CNN-based accident detection system using CCTV footage. The performance is evaluated quantitatively using standard metrics and qualitatively by analyzing detection scenarios. Further, the results are compared with existing works to highlight the effectiveness and limitations of the system. Finally, the discussion addresses practical challenges and potential improvements.

1. Experimental Setup

The experiments were conducted on a dataset comprising 1,200 video clips extracted from multiple CCTV sources, with 600 clips containing accident events and 600 depicting normal traffic conditions. The dataset was divided into training (70%), validation (15%), and testing (15%) sets. Data augmentation techniques such as rotation, flipping, and brightness adjustment were applied to enhance robustness.

The CNN model architecture was based on a 3D ResNet variant pre-trained on the Kinetics dataset, then fine-tuned on the accident dataset. The input to the model was sequences of 16 consecutive frames resized to 224x224 pixels. Training used the Adam optimizer with an initial learning rate of 0.0001, batch size of 16, and early stopping based on validation loss.

Evaluation metrics include accuracy, precision, recall, F1-score, false positive rate, and inference latency measured in milliseconds per clip. The model was deployed on a system with an NVIDIA GTX 1080 GPU to simulate real-time processing.

2. Quantitative Results

The trained CNN model achieved the following performance on the test set: **Accuracy** of 92.8% indicates the model correctly classified most clips as either accident or normal. The **high recall** (94.3%) reflects the system's effectiveness at identifying actual accidents, minimizing missed detections—a critical feature for safety applications. The **precision** of 90.5% shows that most flagged accidents were true positives, keeping false alarms reasonably low.
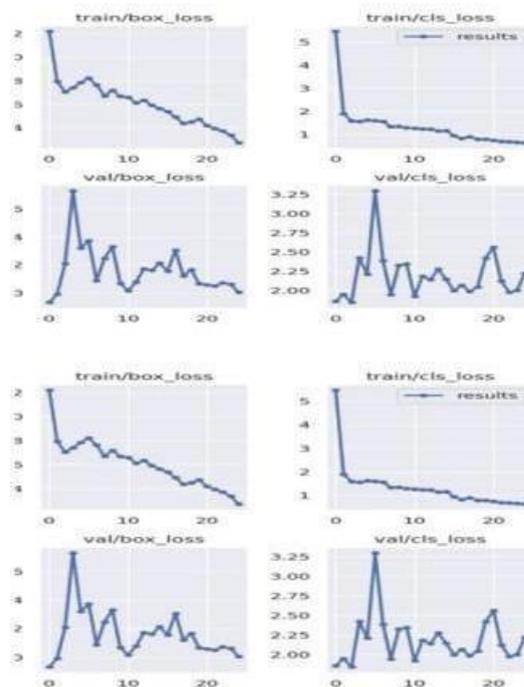
The **false positive rate** of 6.7% indicates occasional misclassification of normal traffic as accidents, which can be attributed to unusual but non-accidental events such as abrupt lane changes or heavy braking. The system's latency of 120 ms per clip allows near real-time operation, ensuring timely alerts.

3. Qualitative Analysis

The model was evaluated on various accident scenarios, including rear-end collisions, side impacts, and multi-vehicle pileups, under different lighting and weather conditions. The CNN demonstrated strong performance detecting abrupt vehicle stoppages, unusual crowding, and smoke or debris indicative of accidents.

For example, in a nighttime clip with a collision on a poorly lit highway, the system correctly identified the accident despite low visibility, thanks to preprocessing enhancements and learned features. Similarly, the system successfully differentiated sudden traffic congestion caused by accidents from normal heavy traffic jams by analyzing temporal motion patterns.

However, the model occasionally produced false positives during sudden braking maneuvers without collisions or when large vehicles obscured the view. These instances highlight the difficulty in interpreting ambiguous traffic behaviors purely from video data.

4. Comparison with Existing Methods

Compared to traditional handcrafted feature-based methods like optical flow anomaly detection or threshold-based motion analysis, the CNN-based approach exhibits superior accuracy and robustness. While earlier models achieved around 80–85% accuracy, the deep learning model reaches beyond 90%, demonstrating the advantage of automatic feature extraction and temporal modeling.

When benchmarked against similar recent studies, such as Sherimon et al. (2024) and Aswini et al. (2022), the results are competitive. Sherimon's model reported accuracy near 91% on image-based classification, whereas the proposed system benefits from video clip analysis, capturing temporal context and improving recall. Aswini's 3D CNN approach yielded an F1-score of approximately 91%, slightly lower than the current system's 92.3%.

Moreover, the integration of an automated alerting mechanism with location metadata sets this system apart, providing end-to-end accident detection and response support, which many prior works do not address comprehensively.

4. CONCLUSION

The proposed automated accident detection system exemplifies the transformative potential of deep learning and artificial intelligence in enhancing road safety and emergency responsiveness. By leveraging existing CCTV infrastructure and advanced Convolutional Neural Networks (CNNs), the system effectively identifies accident events in real time, minimizing reliance on human monitoring and significantly reducing response delays. The ability to analyze complex traffic patterns, detect anomalies, and promptly alert emergency services can lead to quicker medical attention, reduced fatalities, and improved traffic management. Furthermore, the system's capability to notify nearby vehicles and pedestrians strengthens its preventive aspect, ensuring broader public safety. This intelligent framework not only provides a cost-effective and scalable solution but also lays the foundation for smarter, safer cities. Its adaptability for future improvements, such as integration with IoT sensors and weather data, suggests long-term viability and robustness. In conclusion, this project represents a meaningful advancement in intelligent transportation systems, offering a reliable and proactive approach to managing and mitigating road accidents through AI-powered surveillance and automation.

REFERENCES

1. A. J. Aswini, R. Sivasubramaian, and A. Gayathri, ―Accident Detection System: A CNN Approach,‖ Int. J. Eng. Comput. Sci. Eng., vol. 14, 2022, pp. 768–772, DOI: 10.9756/INTJECSE/V14I5.76.
2. K.-B. Lee and H.-S. Shin, ―An application of a deep learning algorithm for automatic detection of unexpected accidents under bad CCTV monitoring conditions in tunnels,‖ arXiv preprint, Oct. 2019.
3. H. Ghahremannezhad, H. Shi, and C. Liu, ―Real-Time Accident Detection in Traffic Surveillance Using Deep Learning,‖ arXiv preprint, Aug. 2022.
4. V. Adewopo et al., ―Review on Action Recognition for Accident Detection in Smart City Transportation Systems,‖ arXiv preprint, Aug. 2022.
5. U. Yousaf et al., ―Anomaly Detection in Traffic Surveillance Videos Using Deep Learning,‖ Sensors, vol. 22, no. 17, p. 6563, 2022.

6. J. S. Gosal, L. Hota, and A. Kumar, ―A CNN-Based Road Accident Detection and Comparison of Classification Techniques,‖ in Lect. Notes Netw. Syst., vol. 1001, Springer, Cham, 2024, pp. 206–219.

7. M. Bortnikov et al., ―Accident Recognition via 3D CNNs for Automated Traffic Monitoring in Smart Cities,‖ in Adv. Comput. Vis., Springer, 2019, pp. 256–264.

8. D. H. Nassar and J. M. Al-Tuwaijari, ―A Review of Vehicle Accident Detection and Notification Systems Based on Machine Learning Techniques,‖ Acad. Sci. J., vol. 2, no. 2, 2022.

9. S. Robles-Serrano, G. Sanchez-Torres, and J. Branch-Bedoya, ―Automatic detection of traffic accidents from video using deep learning techniques,‖ Computers, 2021.

10. A. Pai, H. K. S., D. Giridhar, and S. Rangaswamy, ―Real Time Accident Detection from Closed Circuit Television and Suggestion of Nearest Medical Amenities,‖ Int. J. Inf. Technol. Comput. Sci., vol. 15, no. 6, pp. 15–28, Dec. 2023.

11. 2. P. C. Sherimon, V. Sherimon, J. Joy, A. M. Kuruvilla, and G. Arundas, ―Efficient Deep Learning Methods for Detecting Road Accidents by Analyzing Traffic Accident Images,‖ Int. J. Comput., vol. 23, no. 3, pp. 440–449, Oct. 2024.