An Overview of Different Deep Learning Techniques Used in Road Accident Detection

Vinu Sherimon[1], Sherimon P.C.[2], Alaa Ismaeel[3], Alex Babu[4], Sajina Rose Wilson[5], Sarin Abraham[6], Johnsymol Joy[7]
College of Computing and Information Sciences,
University of Technology and Applied Sciences, Muscat, Sultanate of Oman[1]
Faculty of Computer Studies, Arab Open University, Muscat, Sultanate of Oman[2, 3]
PG Department of Computer Applications and Artificial Intelligence,
Saintgits College of Applied Sciences, Kerala, India[4, 5, 6, 7]

***Abstract*—Every year, numerous lives are tragically lost because of traffic accidents. While many factors may lead to these accidents, one of the most serious issues is the emergency services' delayed response. Often, valuable time is lost due to a lack of information or difficulty determining the location and severity of an accident. To solve this issue, extensive research has been conducted on the creation of effective traffic accident detection and information communication systems. These systems use new technology, such as deep learning algorithms, to spot accidents quickly and correctly and communicate important information to emergency workers. This study provides an overview of current research in this field and identifies similarities among various systems. Based on the review findings, it was found that researchers utilised various techniques, including MLP (Multilayer Perceptron), CNN (Convolutional Neural Network), and models such as DenseNet, Inception V3, LSTM (Long short-term memory), YOLO (You Only Look Once), and RNN (Recurrent Neural Network), among others. Among these models, the MLP model demonstrated high accuracy. However, the Inception V3 model outperformed the others in terms of prediction time, making it particularly well-suited for real-time deployment at the edge and providing end-to-end functionality. The insights gained from this review will help enhance systems for detecting traffic accidents, which will lead to safer roads and fewer casualties. Future research must address several challenges, despite the promising results showcased by the proposed systems. These challenges include low visibility during nighttime conditions, occlusions that hinder accurate detection, variations in traffic patterns, and the absence of comprehensive annotated datasets.**

***Keywords*—Deep learning; road traffic; road accident detection; MLP; CNN; LSTM; DenseNet; RNN; inception V3**

I. Introduction

Road crashes are one of the most prominent reasons for death, disability, and hospitalization of people worldwide. 1.35 million people worldwide lose their lives due to road accidents every year [1]. Almost 3,700 people worldwide lose their lives in collisions with automobiles, buses, motorbikes, bicycles, lorries, or pedestrians every day [1]. Cyclists, motorcyclists, and pedestrians account for more than half of the fatalities. According to estimates, crashes rank as the eighth most common cause of mortality worldwide across all age categories and the most common cause of death for children and young adults (ages 5 to 29) [1]. The traffic authority has seen that many vehicles are travelling at high speeds and without fear in crowded areas. The results of traffic accidents are fairly evident. To address these issues, though, a reliable traffic monitoring and management system is needed [2].

Unfortunately, the age group most affected by road accidents is people aged between 18 and 45 years, which accounts for approximately 70% of accidental deaths [3]. Most of these fatalities occur due to a lack of medical care given at the appropriate time. The expenditures associated with medical care, property damage, legal actions, and reduced worker productivity are significant. In addition to the human cost, road accidents place significant economic pressure on countries, taxing healthcare systems and diverting revenue away from other vital sectors, which hinders economic expansion. Due to its vital impact on public safety, economic stability, and general societal well-being, road accident detection is of the utmost importance. To prevent these serious consequences, effective road accident detection is urgently required.

Current accident reporting systems that rely on reports from witnesses or have slow response times frequently cause delays in facilitating traffic control and providing emergency help. The lengthening of travel times for other road users as well as secondary accidents, congestion, and a worsening of injury severity can all result from this delay.

Deep learning [4] is a subset of machine learning that employs multilayer neural network models that are designed to function similarly to the human brain. Unlike other machine learning approaches, deep learning can automatically learn from data such as photos, videos, or text without the aid of any subject expertise. Deep learning employs a multi-layered neural network to extract features from the data and gets better and better at recognizing and classifying data on its own rather than depending on labels included in the unstructured data. In areas like machine vision, computer vision, speech recognition, audio recognition, natural language processing, social network filtering, pharmaceutical engineering, bioinformatics, medical image analysis, and gaming programs, deep learning architectures like deep neural networks, deep belief networks, recurrent neural networks, and convolutional neural networks have been used with related and, in some cases, improved outcomes.

A subfield of artificial intelligence (AI) called computer vision [5] allows computers and other systems to retrieve useful data from photos, videos, and other visual inputs, to analyze objects in photos and videos just like people do, and to take decisions or offer suggestions based on that data. Computer vision tasks include methods for acquiring, processing, and analyzing digital images to generate numerical or symbolic data. Using digital images and deep learning models, computer vision allows computers to accurately identify and classify objects and respond to them. A convolutional neural network understands single images, whereas recurrent neural networks work with video inputs, allowing computers to learn how a series of images is related to one another. Computer vision software can process images in real-time to detect road edges, read traffic signals, and identify other cars, objects, and pedestrians from the camera output.

Deep learning is revolutionizing road accident detection systems by using neural networks to improve responsiveness, accuracy, and adaptability. Deep learning allows these systems to use convolutional neural networks (CNNs), recurrent neural networks (RNNs), and other sophisticated designs to analyze complex visual and sensor data collected from various sources. Deep learning algorithms may develop robust and adaptable models that can cope with the intricacies of the real world by learning from a variety of datasets that cover a range of weather conditions, lighting conditions, and road types. Additionally, its capacity for real-time processing makes it easier to discover accidents quickly, enabling emergency services and traffic control authorities to act right away.

Several research gaps and unresolved issues emerged as the literature on deep learning-based road accident detection systems continued to develop. Even though deep learning techniques have been used in this field in important ways, a thorough synthesis of the literature in the field reveals some gaps that call for more research. This review article offers a current assessment of the state of the field, considering substantial advancements, new patterns, and cutting-edge technology in road accident detection systems. A current summary of the most recent developments in the field is covered. This research paper advances knowledge in this field by methodically evaluating and synthesizing the body of literature on deep learning-based road accident detection systems. This analysis can help researchers and practitioners and can act as a starting point for new research and technological breakthroughs. Policymakers and emergency service providers can prioritize investments in technology, infrastructure, and staff by comprehending the benefits and drawbacks of various deep learning-based solutions.

The rest of the paper is organized as follows: Section II describes the methodology employed to conduct the literature review and discusses the databases, search terms, and inclusion/exclusion criteria used to select relevant papers. Section III then includes a detailed analysis of each selected piece of literature, discussing the main findings, and contributions of each study. Also, a concise summary of the key findings and outcomes of the reviewed papers is included in this section. Section IV provides conclusion, summarizes the main findings and contributions of the literature review, and Section V provides recommendations for future research directions.

II. Materials and Methods

A. Literature Sources and Search Strategies

The information for this review was obtained by looking at publications that have been published between January 2019 and June 2023 using the well-known search engines Google Scholar, Scopus, and Web of Science. A combination of keywords like artificial intelligence, deep learning, machine learning, artificial neural networks, convolutional neural networks, and accident detection was used for data searching.

B. Identification and Selection of Relevant Studies

The articles from the preliminary search that had abstracts with full texts were retrieved. Along with computer searching, non-electronic sources were also used, such as manual searches for relevant journals and publications. The articles were chosen in accordance with the title and after reviewing the abstracts pertaining to our study subject. 202 publications that fit the review's objectives were primarily found through the initial search. Due to data duplication, nine articles were ignored, and 57 were removed due to topic irrelevance. Thus, 136 articles were included for the second stage of data selection. These articles were then subjected to the subsequent inclusion and exclusion criteria.

C. Eligibility Criteria of the Studies

According to the inclusion criteria, the articles must be specifically focused on deep learning-based road accident detection systems; they must also explicitly address the deep learning technology utilized in the study model; they must explicitly discuss the data sets that were used for the model's evaluation, validation, or training; and they must explicitly mention a quantitatively measurable predictive outcome. Articles written in languages other than English and those with topics other than deep learning technology were excluded. 118 articles were excluded after applying the criteria.

D. Data Extraction and Management

The number of articles was further lowered to 18 after applying these eligibility requirements.

Fig. 1 displays the flowchart of the screening and selection of literature.

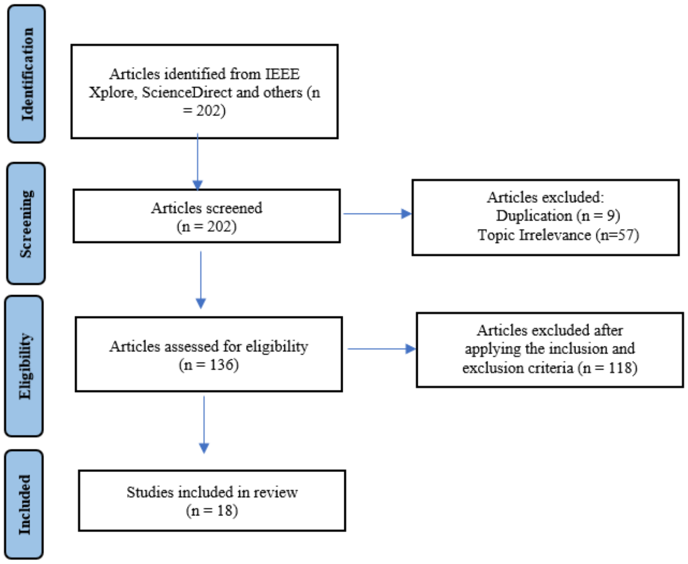

Fig. 1. Flowchart for screening and selection of literature.

III. Discussion

This study in [6] investigates relevant data variables, such as accident severity, the number of casualties, and the number of cars, to uncover trends in road automobile accident data and to suggest a predictive model. Consequently, a pre-processing model is created to transform raw data using interquartile outlier removal, generalization of data attributes, and the elimination of missing and nonsensical features. The performance of road accident prediction is studied using and evaluating four classification methods: naïve Bayes, random forest, decision trees, and multinomial logistic regression. Except for naïve Bayes, the results address acceptable levels of accuracy for car accident prediction. The results are examined using a data-driven methodology to identify the critical variables causing traffic accidents and to suggest ways to prevent them.

In this research [7], a deep learning accident prediction model is proposed by combining extended features such as sentiment analysis, emotions, weather, geocoded locations, and time information with data derived from tweet messages. According to the data, accident detection accuracy increased by 8%, bringing the test's accuracy up to 94%. The suggested methodology fared better than the state-of-the-art methods now in use, increasing accuracy by 2% and 3%, respectively, to reach 97.5% and 90%. The high-performance computing constraints brought forth by detector-based accident detection, which required massive data calculation, were likewise addressed by our method. The outcomes have increased confidence that utilizing cutting-edge features improves traffic accident identification and prediction.

The work done by [8] provides an overview of the technologies used in road accident detection and information transfer system. They also emphasized the importance of automatic road accident detection and information communication systems in all vehicles, whose role is to send a notification message to the nearest hospital and police station upon detecting an accident. As the performance of such systems can be greatly improved with the inclusion of technologies like social network data analysis, sensor data, machine learning, and deep learning, the authors have used artificial neural networks to build a new model. The entire system is structured as three modules, where the first module comprises the collection of image data. The second module is about pre-processing the data and the last module comprises accident detection and notification, which extracts visual and temporal features using the PYTORCH architecture and neural networks called a multi-layer perceptron (MLP). The proposed system proved to be very useful in detecting incidents with an F1 score of 0.98 and accuracy of 98 percent.

The high number of accident-related fatalities in India is a topic of this study [5], which highlights the importance of quick guidance. The paper suggests a method for identifying accidents using real-time CCTV data from roadways. It uses a Hierarchical Recurrent Neural Network (H-RNN) model that has been trained to distinguish between accident- and non-accident-related video frames. H-RNN is a potential option for quick accident detection because of its accuracy in image classification—over 95%—and applicability for video analysis—it can capture temporal dependencies. With the help of this study, accident victims will receive much better response times and outcomes.

The research described in [9] highlighted the need of minimizing the time required by an accident-detection system in sending notifications to the emergency services. To meet this objective, the authors proposed a software-oriented accident recognition system using convolutional neural networks (CNN) that could send real-time alerts to the nearest hospitals and police stations. The new system used only the available resources with the modifications brought through the deployment of AI techniques on edge devices attached to roadside CCTV cameras for classifying images. As the input to the convolutional neural network (CNN), live footage from CCTV cameras were used, from which the probability of the occurrence of an accident was generated. Accident Recognition and Alerting System (ARAS) is divided into three main components, based on their functionality. The components are the Edge Accident Recognition Module (EARM), Alert Database (AD) and Alert Web Interface (AWI). The EARM is responsible for recognizing accidents and sending alert messages. The alerts are received by the AWI, which is a website provided to the emergency services. The AD acted as a common point between the EARM and the AWI. Three kinds of CNN architectures are used which are Inception v3, DenseNet, and ResNet-50 respectively thus resulting in three new models. The performance of these models was compared, and it was found that Inception v3 detected the accident throughout as opposed to DenseNet and ResNet-50. Also, the Inception v3 model displayed the best prediction time, proving it to be an effective choice for real-time, end-to-end, edge deployment. These models additionally help in avoiding the transmission of duplicate alerts.

A new support system for self-driving cars is proposed in [10] that could detect vehicular accidents through a dashboard camera and spontaneously report it to local authorities. The Mask R-CNN framework is used for object detection along with an object tracking algorithm to make the new system more efficient. The object detection component serves to identify the type of vehicle as either a car, bus, or truck. The output of this step includes the class IDs, generated masks, bounding boxes, and the detection scores of vehicles for a video frame. Whereas the part comprising a centroid tracking algorithm is used for vehicle tracking.

Additionally, trajectory intersection, velocity computation, and lane detection are also included to improve the performance of the vehicular accident detection framework. The proposed model is compared with existing models using two important performance criteria, i.e., Accident Detection Rate (ADR) and False Alarm Rate (FAR). As the number of parameters is more, the reliability of the new system in accurately recognizing accidents also got magnified with 79.05% Accident Detection Rate and 34.44% False Alarm Rate. Despite these advancements, the system is handicapped by the unavailability of large, annotated datasets that can be used to test and improve the system further.

The research study in [11] proposed Object Detection and Tracking System (ODTS) for automatically detecting and monitoring unexpected events on CCTVs in tunnels. The unexpected events are categorized as either Wrong-Way Driving (WWD), Stop, Person out of the vehicle in a tunnel, or fire. The system accepts selected frames of video at specified time intervals as input, which are then converted into Bounding Box (BBox) and is simultaneously classified into its corresponding type by the detection module which uses a faster RCNN learning algorithm for the same. The detected objects are then assigned a unique ID number by the object tracking module which uses the SORT framework. However, automatic object detection systems in tunnels suffer from the common problem of low illumination videos.

A deep learning model of Faster R-CNN was used for training dataset to address this issue. However, the model does have a limitation in that the ODTS employs an object tracking function with car objects only. The ODTS module applied the Car Accident Detection Algorithm (CADA) for classifying several events. The inclusion of these technologies resulted in the system being able to identify the incidents within 10 seconds after testing with the image that contained each incident but due to the lack of sufficient Fire objects in the training set, the likelihood of false detection in the case of Fire is high.

An automated system for detection of accidents at traffic intersections is proposed in [12]. The proposed framework is made of three main modules, which are object detection, object tracking and the accident detection module respectively. The object detection part is primarily based on the YOLOv4 approach, which makes it very efficient and accurate. The item monitoring module uses the Kalman filter and the Hungarian algorithm for affiliation whose performance is further enhanced with the use of a new cost function. It can accommodate several challenging situations, such as occlusion, overlapping objects, and shape changes that occur while tracking the objects.

Whereas the accident detection module was constructed with the aid of trajectory conflict analysis. The trajectories of the objects are analysed in phrases of velocity, angle, and distance so that extraordinary forms of trajectory conflicts consisting of vehicle-to-vehicle, vehicle-to-pedestrian, and vehicle-to-bicycle can be detected. These unique features make the model suitable for real-time traffic surveillance. The performance of the proposed framework is evaluated using the video sequences under numerous illumination conditions that were obtained from YouTube. The robustness of the system is proved through the lower fake alarm rates and an excessive detection rate.

A traffic accident prediction system that predicts traffic accidents on highways is proposed in study [13]. To accurately predict accidents, the model uses Java's high-level neural network API framework. Here are the steps the model goes through: A convolutional layer is first given traffic accident records as input and is then filtered to remove duplicate fields. After removing the unwanted records, the adjusted dataset is fed to the new layer. Weights are then assigned to the retrieved features based on random number generation, and finally network backpropagation is performed until the error reaches a threshold. The performance of the proposed convolutional neural network (CNN) is evaluated and compared with the traditional neural network-based backpropagation prediction methods. The results show that the new algorithm has much lower loss and higher prediction accuracy compared to the conventional devices. The incident prediction model proposed in this paper has only a small impact on the prediction rate, but it can be improved by choosing other features to input to the CNN or applying a more efficient model instead of the CNN.

A framework for automatic detection of road accidents in surveillance videos is proposed in [14]. A convolutional auto-encoder was incorporated in the model for extraction of deep representation, in the form of appearance and motion features as well as their correlations, from the spatiotemporal video volumes (STVVs) of raw pixel intensity and then detection for any abnormalities in surveillance videos. The accident detection ability of the model is further increased by combining the complementary appearance and motion information. Additionally, a collision is also sensed using the information about the joints of the trajectories of the vehicles over space-time dimensions. Once the joint trajectories are made, a one-class SVM with an RBF kernel is used to generate the outlier score from his intermediate representation to identify unusual/unseen/abnormal/outlier events. The performance measure of this model is evaluated using real accident videos collected from the

CCTV surveillance network of Hyderabad City in India. The result witnessed a reduction in the false alarm rate and an increase in the reliability of the detection system by using the proposed approach. But the system still faced the challenges of low visibility at night, occlusions, and large variations in the normal traffic pattern, which need to be addressed in the future.

To ensure that victims receive timely aid, the study given in [15] created a system to identify traffic incidents in real time and send an immediate alarm message to the nearest control unit. As the existing structures that use features such as IR sensors, IMU sensors, ARDUINO UNO etc., lack accuracy in detecting accidents, the authors attempted to build in more accuracy and effectiveness in their system through the usage of deep learning techniques such as convolutional neural networks. As no dataset regarding accidents occurring in India was available dataset was created to train the CNN model. The trained system is then included with the cameras, positioned at accident-susceptible areas, to seize the video of the vehicles on the road. Then for video classification, deque was used for performing the rolling prediction averaging. Once an accident is detected, officials are informed about the same by sending them an alert message using the GSM module that's included in conjunction with the camera. The proposed version proved to be more dependable and economical compared to the current structures. It could detect accidents with a high level of accuracy with almost eighty-five percent of successful image prediction. The author similarly proposed incorporating speed-detection cameras as future work, which could make the system more efficient.

Using both video and audio data generated from dashboard cameras, the research in [16] proposed a car crash detection system. Multimodal data is composed of three different types of data, namely, video data, audio features of audio data, and spectrogram images of audio data which were collected from a video sharing platform like YouTube. The data were then standardized and normalized to occupy a standard frame. Deep learning techniques like gated recurrent unit (GRU) and convolutional neural network (CNN) were then used to classify the data as either positive or negative clips. The CNN module consisted of five conv2D layers with ReLu activation function, four batch normalization layers with 0.9 momentum, three max pooling layers, and a global max pooling layer. A weighted average ensemble is used as an ensemble technique for combining the three different classifiers. Finally, the classification performance of the proposed car crash detection system was compared with the existing system, and it was found that the proposed system performed better than its base counterpart, which relied on single modal data (i.e., video or audio data only) to detect possible car-crashes. The proposed model could bring in a drastic improvement in the accuracy of the system. The authors further proposed an enhancement to their model using the inclusion of three-dimensional filters, which may further improve the efficiency more.

An automated system capable of detecting traffic accidents from video is proposed in [17]. A deep learning (DL)-based approach is used, which shows a high performance in accomplishing computer vision tasks that involves complex features relationship. The work assumes that accident events can be represented using data that have visual features occurring in a temporal way. The model hence comprises a feature extraction phase that uses the convolutional method, followed by temporary pattern identification accomplished using recurrent neural networks. For accident detection, a dense ANN approach is used.

Therefore, the model can be structured into three portions where each portion focuses on spatial feature extraction, temporal feature extraction and binary classification, respectively. The model was trained using the imageNet dataset. As the precision of the accident detection system was found to be under satisfactory, the model was modified using a transfer learning process and the weights for the model were adjusted using a new dataset. The modification included a ConvLSTM layer-based neural network being added up to the existing architecture for enhanced feature extraction and usage of a dense artificial neural network block using regularization methods for better accident detection.

Under the ConvLSTM layer-based neural network model, the feature vector computed by the adjusted InceptionV4 architecture was taken as the input. For this, two sets of data are used, where the first one comprised image for the visual extractor and the other had videos for training the temporal extractor. Both data sets contain positive and negative accidents. The modification also included the application of four techniques in video segmentation, in which the first one consisted of video segmentation without frame discrimination., the second one was used to skip the frame intruder to reduce the redundancy, the third technique helped in calculating a pixel-to-pixel comparison of two consecutive images and finally the fourth technique is substantially similar to the third except that the threshold was set to 0.98 against the 0.9 mark in the third. Through the modification the accuracy improved to 0.98. But the proposed model still had a few limitations, which is, the model required large datasets with clear data to get trained well.

A feature fusion-based deep learning model for video-based accident detection is suggested in study [18], that could achieve better detection speed and accuracy with restricted computing resources. The system comprises an interest module for capturing the appearance features of the crash images. The module was blended with ResNet to enhance the speed of the traditional convolution neural network. A 1 x 1 convolutional layer is employed to reduce the size of the output feature map, which is then entered into the Conv-LSTM network to accurately extract the motion features of crashes. This Conv-LSTM network has a bonus over its traditional counterpart in terms of being lighter and preserving spatial information. Finally, a global pooling layer is used to locate a crash. This proposed system won over the current video-based crash detection systems as they suffered from low detection accuracy and excessive computational costs. However, the proposed version has a few limitations like falsely detecting traffic scenes as crashes, showcasing few misdetections in complicated, rare, and ambiguous environment, etc. The model may also be, in addition, stepped forward to become aware of one-of-a-type types/severity tiers of crashes.

An accident detection system that may track accidents at the generation of occurrence is given in [19]. The proposed model is a fusion of CNN and LSTM layers for the classification of nonstop video captured by the camera. In the CNN-LSTM network, CNNs are used for image feature extraction and exceed LSTMs for sequence prediction. When compared with other existing models composed of high-priced sensors and pointless hardware, the proposed model showed improved performance in terms of cost, reliability, and accuracy. In the future, the model's performance may be enhanced using some supervised and unsupervised techniques. Particularly, supervised learning can be used to discover the accidents from the frames which can be flagged anomalously via way of means of unsupervised models. More state-of-the-art techniques for monitoring accident-prone areas may be used.

A system to detect accidents from video footage provided to it using a camera is proposed in [20]. The objective of this model, which is built using advanced Deep Learning Algorithms, i.e., convolutional neural networks (CNN) and the LSTM network is to detect accidents within seconds of its occurrence. In the CNN-LSTM network, CNN is used for modeling spatial data like images and feature extraction which is passed on to the LSTM for sequence prediction. The pictures captured through the Pi-cameras are broken down into frames and fed to the automated system and while an accident is detected, an alert message is dispatched using the GSM module. Along with the message, the frame at which an accident is detected and the percentage of chances of accident are also conveyed. Upon assessing the model's performance, it yields an average accuracy of 92.83%.

An automatic car accident detection method based on Cooperative Vehicle Infrastructure Systems (CVIS) and machine vision such that roadside intelligent devices recognize and locate crashes efficiently is given in [21]. For this purpose, a new image dataset (CAD-CVIS) is created based on video sharing websites consisting of various kinds of accident types, weather conditions, and accident locations. The data includes video features of the car's motion parameters, and the concept of accident detection is since the motion parameters will change dramatically when an accident occurs. The CAD-CVIS dataset used in the study was divided into three parts. A training set (80%) used to train the parameter weights of the network, a validation set (5%) used to fir and test hyperparameters such as learning rate and dropout rate and a test set (15%) used to evaluate the performance of various algorithms for detecting car accidents. Here, each part of the dataset contains all types of accident.

A deep neural network algorithm, YOLO-CA, was used to detect accidents and their location. The performance of the model was enhanced by including Multi-Scale Feature Fusion (MSFF) and a dynamically weighted loss function to help detect small objects better. Finally, the performance of the proposed model was evaluated and compared with other object detection models, which showed significant improvements in terms of response time by rescue agencies and rescue efficiency.

A prototype that uses a machine learning approach to reduce road accidents caused by reasons such as drowsiness, fatigue, and inattention is developed in [22]. The CNN algorithms help avoid accidents, notify drivers when they detect drowsiness, and are very efficient at classifying datasets. This approach combines machine learning techniques with other concepts such as avoidance of drinking, direction control, speed control, and distance maintenance. The new system is superior to many existing systems as it uses more sophisticated and robust sensors than the others. Additionally, the proposed system uses IoT components cheaper than available mechanisms such as Wireless Sensor Networks (WSN), Cloud computing, and Industry 4.0, etc., so it offers a high-tech solution to avoid accidents at low cost. In conclusion, it can be stated that the proposed system offers a great opportunity to avoid accidents at an affordable price.

Table I presents an overview of the reviewed papers. A comprehensive overview of the reviewed papers, highlighting the dataset used, technique employed, observations made, and performance, is given in the table. The dataset used column provides insight into the specific data sources utilised in the studies. This information is essential for understanding the scope and applicability of the research findings. The methodology column outlines the methodologies, algorithms, or experimental approaches employed by the authors. This section sheds light on the scientific tools and methods utilised in each study, allowing readers to assess the validity of the research. The observation column summarises the key findings or outcomes of each paper. The performance column summarises the accuracy of the models.

TABLE I. SUMMARY

Author & Year	Dataset used	Methodology	Observations	Performance
(Pourroostaei Ardakani et al., 2022)	Online traffic accident dataset based on UK	Decision trees, Random Forest, Multinomial logistic regression, and Naïve Bayes algorithm	This paper presents a predictive model for road car accidents, focusing on significant data features like accident severity, casualties, and vehicles. Four classification methods are used, with acceptable accuracy levels. Among the four, random forest algorithm can predict car accidents with more accuracy.	85% accuracy
(Azhar et al., 2022)	Twitter messages	Deep learning techniques	The paper presents a deep learning accident prediction model that uses tweet messages and additional features like sentiment	94% accuracy

			analysis, emotions, weather, and geo-coded locations.	
(L et al., 2022)	The web scraping method was used to construct the image dataset from scratch.	Feed forward Neural networks (MLP)	Visual and temporal feature extractors are used to distinguish between traffic collisions. The training process for fine-tuning the weights used the picture dataset. Because there aren't many instances given, the solution only applies to car accidents—motorcycles, bicycles, and pedestrians aren't included.	92% accuracy
(Parthiban et al., 2021)	CADP dataset, which includes recordings with accidents, and the DETRAC dataset, originally designed for vehicle object location.	Hierarchical Recurrent Neural Network	Compared to traditional RNNs, the H-RNN is more appropriate for video extraction. H-RNN based picture classifiers have provided precisions of over 95% for comparatively smaller datasets and require less preparation.	92.38% accuracy
(Chitale et al., 2020)	The TrafficNet dataset from DeepQuest AI	CNN, InceptionV3, Densenet and Resnet-50	A software-oriented approach employing CNN to detect road accidents. This produces the likelihood that an accident may occur or not. In contrast to DenseNet and ResNet-50, Inception v3 identified the accident continuously. The best prediction time was also achieved by the Intermediate Representation of Inception v3, indicating that this model is a good option for real-time, end-to-end edge deployment.	93.75% accuracy
(Chand et al., 2020)	The Microsoft Common Objects in Context dataset and dash-cam footage of accidents from YouTube	Mask R-CNN	The proposed model compared Accident Detection Rate (ADR) and False Alarm Rate (FAR). The suggested approach has an accurate accident detection rate of 79.05% for cars and a false alarm rate of 34.44% for cars.	79.05% accuracy
(Lee & Shin, 2019)	CCTV footage	Faster RCNN	Proposed Object detection and tracking system (ODTS) by fusing an object tracking method with a deep learning-based object detection procedure. The model achieved average precision values of 0.8479, 0.7161, and 0.9085 for cars, people, and fires, enabling the system to detect all accidents in less than 10 seconds.	84.79% precision for detecting cars
(Ghahremannezhad et al., 2022)	Customized dataset	YOLOv4 method with a pre-trained CNN model CSPDarknet53.	The Euclidean distances between all object pairs are calculated to identify the objects. The angle of collision trajectory conflicts involves near-accident and accident occurrences. With a false alarm rate of 6.89% and a detection rate of 93.10%, the proposed structure performed well.	93.10% detection rate
(Thaduri et al., 2021)	Traffic accident dataset	CNN, backpropagation	Polling method removes unnecessary data. The proposed CNN is compared with the backpropagation prediction method. Compared to the conventional BP algorithm, the suggested algorithm has a substantially smaller loss and a higher prediction accuracy.	98% accuracy
(Singh & Mohan, 2019)	CCTV Surveillance footage	Autoencoder, one class SVM	The proposed framework uses denoising autoencoders trained on traffic videos to automatically learn feature representation from spatiotemporal volumes of raw pixel intensity, replacing traditional hand-crafted features.	77.5% detection rate
(Rajesh et al., 2020)	Customized dataset	CNN	Dropout is used to prevent overfitting. The model was created using sequential API GSM module for alert. The proposed system uses a GSM module to detect road accidents, sending alert messages to nearby control rooms. It is reliable, economical, and highly accurate.	85% accuracy
(Choi et al., 2021)	Images from dashboard cameras	GRU, CNN	Utilizing video data, a CNN-and-GRU-based classifier, a GRU-based classifier utilizing audio characteristics from audio data, and a CNN-based classifier using spectrogram images from audio data make up the suggested model.	98.60% accuracy for Case Study 1 and 89.86% for Case Study 2
(Robles-Serrano et al., 2021)	Customized dataset	ConvLSTM	The suggested approach assumes that visual characteristics that appear over time describe traffic accident incidents. Thus, the model architecture consists of an extraction phase for visual features and a transient pattern identification phase.	98% accuracy
(Lu et al., 2020)	Traffic image datasets	ConvLSTM, ResNet, Vgg	The framework proposes a residual neural network and attention modules for extracting crash-related appearance features from urban traffic videos, which are then combined with Conv-LSTM for simultaneous capture.	87.78% accuracy
(T.S. et al., 2021)	CCTV image dataset	CNN, LSTM	The proposed model combines CNN and LSTM layers for continuous video classification, incorporating ResNet-50 and LSTM layers for spatial and temporal characteristics, with adjustments made for training images.	99.9% accuracy
(Ghosh et al., 2019)	CCTV video dataset	CNN, RNN	The proposed model combines CNN and LSTM layers for continuous video classification, inspired by Inception v3, with temporal and spatial features added to the existing Convolution	92.38% accuracy

			Network, divided into convolution and recurrent parts.	
(Tian et al., 2019)	The novel image dataset CAD-CVIS.	YOLO-CA	The paper presents an automatic car accident detection method using Cooperative Vehicle Infrastructure Systems (CVIS) and machine vision, utilizing a novel image dataset and a deep neural network model YOLO-CA, which enhances performance in detecting small objects.	90.02% average precision
(Razeeth et al., 2021)	Customized image datasets	CNN	The study uses machine learning to detect drowsiness, improve drunk and speed control, and prevent collisions using IoT. It uses convolutional neural network with Keras and MQ13, MAX30105, and L298N sensors.	Drowsiness is detected successfully, and collisions are prevented

IV. Conclusion

The analysis of related work highlights several studies and their methodologies. For instance, researchers have developed models using artificial neural networks, convolutional neural networks (CNN), and object detection algorithms to improve road accident detection. These models utilize image data, pre-processing techniques, and accident detection and notification modules to achieve high accuracy and efficiency. Techniques like social network data analysis, sensor data, and trajectory conflict analysis have also been employed to enhance detection capabilities. Researchers employed a variety of techniques like MLP, CNN, and CNN models, including DenseNet, Inception V3, LSTM, YOLO, RNN, and others. Although the MLP model showed high accuracy, the Inception v3 model provided the best prediction time, proving its suitability for real-time, end-to-end edge deployment.

The proposed systems demonstrate promising results, including high accuracy rates, low false alarm rates, and efficient real-time detection. However, challenges such as low visibility at night, occlusions, variations in traffic patterns, and the lack of annotated datasets still need to be addressed in future research. In most of the papers, the delays in giving timely medical attention to the injured are quoted as the major reason for the increase in traffic fatalities. The models used both video and audio data from the dashboard camera for detection. The major problem with these types of datasets is their size. To develop effective deep learning models, we must train our models on a large dataset.

In conclusion, this review reveals significant advancements in road accident detection using machine learning and deep learning approaches. The studies highlight the importance of accurate and efficient detection systems that can improve emergency response and reduce the impact of accidents.

V. Future Work

To overcome the problems related to data scarcity, we are planning to use different strategies, like different pre-trained convolutional neural networks or Spinal net, for feature extraction. Also, further research is needed to overcome the identified challenges in the existing literature and enhance the reliability and performance of such systems.

Acknowledgment

The research leading to these results has received funding from the Research Council (TRC) of the Sultanate of Oman under the Block Funding Program BFP/RGP/ICT/21/362.

References

[1] CDC, "Road Traffic Injuries and Deaths—A Global Problem," Centers for Disease Control and Prevention. Accessed: Nov. 20, 2023. [Online]. Available: https://www.cdc.gov/injury/features/global-road-safety/index.html

[2] Y. Golhar and M. Kshirsagar, "Emerging Technologies for Driving Road Safety and Traffic Management for Urban Area," J. Comput. Sci., vol. 17, no. 11, pp. 1104–1115, Nov. 2021, doi: 10.3844/jcssp.2021.1104.1115.

[3] "Road Accidents in India 2021." Accessed: Nov. 20, 2023. [Online]. Available: https://morth.nic.in/sites/default/files/RA_2021_Compressed.pdf

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.

[5] D. Chand, S. Gupta, and I. Kavati, "Computer Vision based Accident Detection for Autonomous Vehicles." arXiv, Dec. 20, 2020. doi: 10.48550/arXiv.2012.10870.

[6] S. Pourroostaei Ardakani et al., "Road Car Accident Prediction Using a Machine-Learning-Enabled Data Analysis," Sustainability, vol. 15, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/su15075939.

[7] A. Azhar et al., "Detection and prediction of traffic accidents using deep learning techniques," Clust. Comput., vol. 26, no. 1, pp. 477–493, Feb. 2023, doi: 10.1007/s10586-021-03502-1.

[8] L. Sandeep et al., "A Review on Accident Detection Using Deep Learning to Reduce the Response Time for Medical Help", Int. Res. J. Mod. Eng. Technol. Sci., vol. 04, no. 05, pp. 1907 - 1912, May. 2022

[9] P. Chitale, T. Dhope, A. Akhelikar, and S. Dholay, "Smart Accident Recognition and Alerting System for Edge Devices," in 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Oct. 2020, pp. 486–491. doi: 10.1109/ICCCA49541.2020.9250902.

[10] D. Chand, S. Gupta, and I. Kavati, "Computer Vision based Accident Detection for Autonomous Vehicles," in 2020 IEEE 17th India Council International Conference (INDICON), Dec. 2020, pp. 1–6. doi: 10.1109/INDICON49873.2020.9342226.

[11] 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," in 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey: IEEE, Aug. 2019, pp. 7–11. doi: 10.1109/Deep-ML.2019.00010.

[12] H. Ghahremannezhad, H. Shi, and C. Liu, "Real-Time Accident Detection in Traffic Surveillance Using Deep Learning," in 2022 IEEE International Conference on Imaging Systems and Techniques (IST), Jun. 2022, pp. 1–6. doi: 10.1109/IST55454.2022.9827736.

[13] A. Thaduri, V. Polepally, and S. Vodithala, "Traffic Accident Prediction based on CNN Model," in 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), May 2021, pp. 1590–1594. doi: 10.1109/ICICCS51141.2021.9432224.

[14] D. Singh and C. K. Mohan, “Deep Spatio-Temporal Representation for Detection of Road Accidents Using Stacked Autoencoder,” IEEE Trans. Intell. Transp. Syst., vol. 20, no. 3, pp. 879–887, Mar. 2019, doi: 10.1109/TITS.2018.2835308.

[15] G. Rajesh, A. R. Benny, A. Harikrishnan, J. Jacob Abraham, and N. P. John, “A Deep Learning based Accident Detection System,” in 2020 International Conference on Communication and Signal Processing (ICCSP), Jul. 2020, pp. 1322–1325. doi: 10.1109/ICCSP48568.2020.9182224.

[16] J. G. Choi, C. W. Kong, G. Kim, and S. Lim, “Car crash detection using ensemble deep learning and multimodal data from dashboard cameras,” Expert Syst. Appl., vol. 183, p. 115400, Nov. 2021, doi: 10.1016/j.eswa.2021.115400.

[17] S. Robles-Serrano, G. Sanchez-Torres, and J. Branch-Bedoya, “Automatic Detection of Traffic Accidents from Video Using Deep Learning Techniques,” Computers, vol. 10, no. 11, Art. no. 11, Nov. 2021, doi: 10.3390/computers10110148.

[18] “A New Video-Based Crash Detection Method: Balancing Speed and Accuracy Using a Feature Fusion Deep Learning Framework.” Accessed: Nov. 08, 2023. [Online]. Available: https://www.hindawi.com/journals/jat/2020/8848874/

[19] “IJISRT21JUN878.pdf.” Accessed: Nov. 08, 2023. [Online]. Available: https://ijisrt.com/assets/upload/files/IJISRT21JUN878.pdf

[20] S. Ghosh, S. J. Sunny, and R. Roney, “Accident Detection Using Convolutional Neural Networks,” in 2019 International Conference on Data Science and Communication (IconDSC), Mar. 2019, pp. 1–6. doi: 10.1109/IconDSC.2019.8816881.

[21] “Tian et al. - 2019 - An Automatic Car Accident Detection Method Based o.pdf.” Accessed: Nov. 08, 2023. [Online]. Available: https://ieeexplore.ieee.org/ielx7/6287639/8600701/08832160.pdf?tp=&arnumber=8832160&isnumber=8600701&ref=aHR0cHM6Ly9zY2hvbGFyLmdvb2dsZS5jb20v

[22] “Accident Mitigation System with Drowsiness Detection: A Machine Learning and Iot with Hybrid Approach | IEEE Conference Publication | IEEE Xplore.” Accessed: Nov. 08, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9491646