Real Time Accident Detection from Closed Circuit Television and Suggestion of Nearest Medical Amenities

Akanksha A. Pai
Department of Computer Science and Engineering, RV College of Engineering, Bengaluru, India
E-mail: akankshaapai.cs20@rvce.edu.in

Harini K. S.
Department of Computer Science and Engineering, RV College of Engineering, Bengaluru, India
E-mail: hariniks.cs20@rvce.edu.in

Deeptha Giridhar
Department of Computer Science and Engineering, RV College of Engineering, Bengaluru, India
E-mail: deepthag.cs20@rvce.edu.in

Shanta Rangaswamy*
Department of Computer Science and Engineering, RV College of Engineering, Bengaluru, India
E-mail: shantharangaswamy@rvce.edu.in
ORCID iD: https://orcid.org/0000-0002-7559-1912
*Corresponding author

Abstract: The prevalence of automobile accidents as a major cause of violent deaths around the world has prompted researchers to develop an automated method for detecting them. The effectiveness of medical response to accident scenes and the chances of survival are influenced by the human element, underscoring the need for an automated system. With the widespread use of video surveillance and advanced traffic systems, researchers have proposed a model to automatically detect traffic accidents on video. The proposed approach assumes that visual elements occurring in a temporal sequence correspond to traffic accidents. The model architecture consists of two phases: visual feature extraction and temporal pattern detection. Convolution and recurrent layers are employed during training to learn visual and temporal features from scratch as well as from publicly available datasets. The proposed accident detection and alerting system using Convolution Neural Network models with Rectified Linear Unit and Softmax activation functions is an effective tool for detecting different types of accidents in real-time. The system of accident detection, integrated with the alerting mechanism for prompt medical assistance achieved high accuracy and recall rates.

Index Terms: Accident Detection, Accuracy, Convolutional Neural Network, Traffic Management.

1. Introduction

The rapid development of infrastructure and technology has undoubtedly made our lives easier, but it has also brought forth new challenges, especially concerning road accidents and traffic risks. Inadequate emergency facilities and poor response times to accidents are major contributing factors to the high death rate and numerous traffic fatalities in our country. To address this critical issue, Intelligent Transportation Systems (ITS) have emerged as a promising solution to strengthen travel safety and improve transportation systems' overall efficiency.

Existing research has explored various accident detection and warning systems, each proposing innovative approaches to tackle this problem. These works have primarily focused on computer vision techniques, leveraging CCTV footage to detect accidents in real-time. However, some of these models lack a real-time alerting system, leading to delays in providing immediate assistance to the injured. Additionally, the accuracy and scalability of existing systems have been areas of concern.

In this paper, we present a novel approach to accident detection and warning systems that builds upon the existing research while addressing its limitations. Our proposed solution utilizes frames from live CCTV footage for real-time accident detection, capitalizing on the already installed infrastructure without the need for additional sensors. By pre-training a Convolutional Neural Network (CNN) model with a dataset containing frames labeled as accident and non-accident, our system can efficiently detect accidents in real-time.

To improve the accuracy and reliability of our approach, we employ an attention mechanism with CNN, enabling the model to focus on relevant features during accident detection. This enhancement ensures that the model can identify accidents with greater precision, reducing false alarms and improving overall system performance.

Moreover, our system incorporates a real-time alerting mechanism, which promptly notifies emergency services, such as nearby hospitals, about the accident location. The inclusion of latitude and longitude information in the alert is a novel feature in our country, enabling faster dispatch of ambulance services and reducing the time it takes for injured crash victims to receive medical attention.

In terms of implementation, we plan to follow a three-phase approach. In the first phase, we will pre-process the training data, extracting relevant features from the CCTV footage. The second phase involves training the CNN model with the labeled dataset and implementing the attention mechanism to enhance accuracy. Finally, in the third phase, we will integrate the real-time alerting system with the model and deploy the web application, making it accessible to relevant stakeholders and emergency services.

By combining real-time accident detection, attention-based CNN models, and immediate alerting, our proposed system aims to significantly improve travel safety, reduce response time to accidents, and save lives. This paper presents a comprehensive overview of our approach, its advantages over existing systems, and the implementation plan to demonstrate its feasibility and effectiveness in real-world scenarios.

2. Literature Survey

The paper discusses the challenges of detecting accidents using CCTV cameras, such as the need for high-quality video footage and the difficulty of distinguishing between accidents and other events [1]. The authors propose a solution that uses a combination of image processing techniques and machine learning algorithms to detect accidents. The proposed system is able to detect accidents in real time and send signals to emergency services. The paper evaluates the performance of the proposed system on a dataset of real-world traffic videos. The results show that the system is able to detect accidents with high accuracy. The authors also discuss the limitations of the proposed system and suggest future research directions.

The authors developed a real-time computer vision based approach to detect and classify traffic incidents [2]. The article proposes a new computer vision-based system that can detect traffic incidents in both sparse and dense traffic flow environments, while considering low-visibility and varying environmental weather conditions. The system consists of five phases, with the first model using YOLOv5 algorithm and DeepSORT tracker to detect and track vehicles in real-time, achieving a mAP rate of 99.2%. The second model is the core of the system, using the YOLOv5 algorithm to efficiently detect and classify the severity level of accidents, achieving a mAP of 83.3%. The third model aims to classify fire occurrences after an incident using transfer learning algorithm ResNet152, achieving an accuracy rate of 98.955%. Overall, this neoteric computer vision-based system shows promising potential for improving traffic incident detection and response, particularly in challenging conditions.

The authors used deep learning for road accident detection from CCTV footage [3]. The suggested framework consists of three phases: pre-processing, training, and testing. The training phase involves both frame-level and feature-level extractions and uses 30 videos for training and 12 videos for testing. The study compares three different architectures: CNN-RNN, CNN-LSTM, and CNN-GRU. The results indicate that CNN-GRU performs better than the other two architectures, achieving higher accuracy.

The paper reviewed real-time traffic accident detection systems using CCTV. The paper "A Review of Real-Time Traffic Accident Detection Systems using CCTV" provides a comprehensive overview of various real-time traffic accident detection systems based on CCTV cameras [4]. The authors discuss the importance of real-time accident detection in reducing the response time to accidents, providing immediate medical assistance, and improving traffic flow. The paper presents a detailed review of various computer vision techniques such as object detection, motion analysis, and deep learning algorithms used in accident detection systems. The authors also discuss the challenges associated with the deployment of such systems, including privacy concerns, data storage, and computational resources. The paper compares various accident detection systems based on their accuracy, efficiency, and scalability. The authors also present a comparison of the different datasets used for training and testing the accident detection systems. The paper concludes that real-time traffic accident detection systems using CCTV cameras have the potential to reduce the response time to accidents, improve road safety, and save lives. The authors suggest that future research should focus on addressing the challenges associated with the deployment of such systems and developing more efficient and accurate accident detection algorithms.

The paper proposes a system aimed at reducing road accidents, especially among two-wheelers, by providing timely medical aid[5]. The system includes an accelerometer attached to the vehicle, which detects the tilt of the vehicle, and a heartbeat sensor on the user's body, which senses abnormal heartbeats to assess the seriousness of the accident.

Based on this data, the system makes a decision and sends information to a smartphone connected to the accelerometer via GSM and GPS modules. When an accident occurs, the system sends text messages to the nearest medical center and friends, providing the exact location of the accident to save time. The paper emphasizes the importance of reducing accident rates through automatic accident detection and alert systems. It highlights the usage of wireless communication techniques, such as GSM and GPS modules, to quickly notify relevant parties about accidents.

The proposed system integrates Arduino as the central controller, enabling the transfer of information between various components. The vibration sensor serves as a critical component in detecting accidents. By using this system, accidents can be promptly reported and medical assistance can be dispatched, potentially saving lives and mitigating the impact of accidents on roads. Overall, the paper introduces a practical and efficient solution for accident detection and alerts, leveraging technology to enhance road safety and emergency response.

The authors presented a real-time accident detection system using CCTV cameras. The paper "Real-Time Accident Detection System Using CCTV Cameras" proposes a system to detect traffic accidents in real-time using CCTV cameras [6]. The system uses computer vision techniques such as object detection and motion analysis to detect accidents based on changes in the environment. The authors discuss the importance of real-time accident detection in reducing response time and providing immediate medical assistance to accident victims. The system is designed to detect various types of accidents such as collisions, rollovers, and skids. Once an accident is detected, the system generates an alert and sends the location information to the nearest medical facility. The paper presents a detailed description of the proposed system architecture, including the hardware and software components. The authors evaluate the system's performance using a dataset of real-time traffic accidents and report high precision and recall values. The proposed system has the potential to reduce response time to traffic accidents, improve road safety, and save lives. The authors suggest that future research should focus on optimizing the system's performance and improving its scalability for large-scale deployment.

The paper discussed accident detection using convolutional neural networks [7]. The paper proposes a method that involves taking each frame of a video and using a convolutional neural network model that has been trained to classify frames into accident or non-accident. Convolutional neural networks learn visual attributes using small input data squares while maintaining the relationship between the image's pixels. The pooling layers are used to reduce the number of parameters when the provided image is too large. The fully connected layer converts the grid into a vector, and it is added to a layer that is similar to a basic neural network. Inception v3 is used as a benchmark for image classification, which uses a variety of convolutions compared to conventional neural networks. This allows the model to explore the image more thoroughly and extract additional features.

The paper presented an accident detection system using surveillance cameras [8]. The study proposes a method for using extraction techniques on surveillance footage to simplify human life on the road. The approach involves the use of various algorithms and processes to achieve this goal. The suggested method involves creating positive or negative bags of video instances from the dataset for training purposes. The bags are then ranked based on the video's features to determine whether they are anomalous or common occurrences.

The authors proposed a computer vision-based accident detection in traffic surveillance [9]. The paper presents a novel framework for detecting road accidents using modern techniques. The proposed framework utilises Mask R-CNN for precise detection of objects followed by an effective centroid-based object tracking algorithm for surveillance footage. The likelihood of an accident is calculated by examining speed and trajectory abnormalities of a vehicle after coming in contact with other vehicles. The framework is capable of achieving high accuracy in detecting accidents and minimizing false alarms on standard road traffic CCTV surveillance footage. The system achieved a detection rate of 71% and a false alarm rate of only 0.53%.

The research proposes a novel and effective framework for accident detection, which consists of three hierarchical steps [10]. Firstly, it utilizes the state-of-the-art YOLOv4 approach to accurately and efficiently identify objects. Secondly, it applies Kalman filters and the Hungarian algorithm for association to track the objects. Finally, it conducts trajectory conflict analysis to detect accidents by analyzing the motion behaviors of the tracked objects and identifying dangerous abnormalities that could lead to minor or major crashes. The framework employs a deep convolutional neural network model for object identification, and the Kalman filter method for tracking the motion of the identified objects and monitoring their trajectories. The discovered bounding boxes are associated between frames using the Hungarian technique.

The paper described automatic accident detection techniques using CCTV surveillance videos [11]. This paper addresses the issue of road accidents and their impact on traffic safety. Various methods for detecting traffic accidents in real-time using surveillance cameras have been discussed. However, these methods are not entirely effective in addressing all types of accident incidents due to factors such as changing lighting conditions, weather patterns, and traffic patterns. The authors have reviewed and compared several similar strategies in this publication.

The authors developed a real-time traffic accident detection and medical assistance system based on computer vision [12]. Real-time Traffic Accident Detection and Medical Assistance System Based on Computer Vision is a paper that proposes a system to detect traffic accidents and suggest the nearest medical amenities to the accident location. The proposed system uses computer vision techniques to identify accidents in real-time using CCTV cameras installed on roads. The system analyzes the video feed from the CCTV cameras and detects accidents based on motion analysis and object detection. Once an accident is detected, the system generates an alert and sends the location information to the nearest medical facility, which can provide immediate medical assistance to the accident victims. The system also

suggests alternate routes to divert traffic from the accident location to prevent further accidents. The proposed system is evaluated on a dataset of real-time traffic accidents, and the results show that the system can accurately detect accidents and suggest the nearest medical amenities with high precision and recall values. The system can potentially reduce response time to traffic accidents and save lives by providing immediate medical assistance.

M. D. Kulkarni, A. D. Kulkarni, & A. K. Shrivastava (2021) presented an accident detection and severity analysis using deep learning [13]. The paper proposes a comparison between real-time accident detection systems using deep learning and CNN versus VGG16. The system employs Vehicular Ad Hoc Network (VANET) and Internet of Things (IoT) to detect accidents and determine their severity using mechanical and medical sensors installed in vehicles. The hospitals are alerted accordingly. The CNN model is trained with a larger epoch to obtain a more accurate model. The model also uses hidden layers and a loss function to eliminate noise in the image. On the other hand, VGG16 is used for image processing and classification. The images are trained using a pre-existing model, and in case an accident is detected, an alert is sent to the authorities.

The paper proposed an intelligent crash recognition method based on 1DResNet-SVM with distributed vibration sensors [14]. The article proposes a pattern recognition method for identifying highway vehicle collisions with guardrails using distributed vibration sensors (DVS) signals. The method utilizes a 1DResNet and SVM model to extract distinguishable characteristics of the DVS signals, achieving an average recognition rate of 96.2%. Field tests were carried out on three typical events in highway accidents, resulting in a recognition precision of 100% for crashes. The proposed method shows promising potential for improving highway safety and reducing accident risks.

The authors conducted a survey of real-time accident detection and response systems [15]. The paper "A Survey of Real-Time Accident Detection and Response Systems" presents a survey of various real-time accident detection and response systems. The authors discuss the importance of real-time accident detection and response in reducing response time and providing immediate medical assistance to accident victims. The paper provides a comprehensive review of various computer vision techniques such as object detection, motion analysis, and deep learning algorithms used in accident detection systems. The authors also discuss the use of different sensors such as cameras, GPS, and accelerometers in accident detection systems. The paper compares various accident detection and response systems based on their accuracy, efficiency, and scalability. The authors also discuss the challenges associated with the deployment of such systems, including privacy concerns, data storage, and computational resources. The paper concludes that real-time accident detection and response systems have the potential to reduce response time to traffic accidents, improve road safety, and save lives. The authors suggest that future research should focus on addressing the challenges associated with the deployment of such systems and developing more efficient and accurate accident detection algorithms.

2.1. Problem Definition

Numerous accidents take place on the roads of India every day. Timely help and action can help save many lives. Many deep learning and machine learning models have been proposed to study the patterns of accident flow, their trajectories and the detection of accidents.

It is the need of the hour to come up with an efficient model which accurately detects an accident and notifies the nearby hospitals for help. The main aim of our project is to detect accidents via the video inputs from CCTV, process them on our trained CNN model, use the CCTV's IP address to determine its location and send a message to all nearby hospitals to make sure the victims get immediate help.

2.2. Basic Theory and Mathematical Model

The proposed methodology aims to detect road accidents in real-time using Convolutional Neural Networks (CNNs) with an attention mechanism. The CNN model is trained on labeled CCTV footage with cross-entropy as the loss function. By setting a threshold, the system alerts nearby hospitals when the accident probability exceeds it. Evaluation metrics like accuracy and confusion matrix assess the model's performance. Future enhancements involve advanced computer vision techniques, multi-modal data integration, and transfer learning to improve accuracy and road safety. Such a system could potentially save lives by enabling prompt medical assistance and addressing the issue of inadequate emergency facilities on Indian roads.

3. Methodology

The proposed methodology leverages real-time accident detection using Convolutional Neural Networks (CNN) trained on labelled CCTV footage. By integrating an attention mechanism, the model can focus on relevant features, leading to superior accuracy in accident detection. The system's ability to process live video feeds enables prompt identification of accidents without the need for additional sensors, ensuring timely response and assistance for injured victims. The inclusion of a real-time alerting mechanism notifies nearby hospitals with the accident's precise location, a novel feature in the country. This feature significantly reduces the time it takes for medical personnel to reach the accident scene, resulting in a higher likelihood of survival for crash victims. By combining these elements, the proposed methodology enhances overall travel safety, lowers traffic fatalities, and efficiently addresses the problem of inadequate emergency facilities.

3.1. System Requirements

The CCTV cameras are fixed at regular intervals along a road continuously, to detect accidents immediately from the video stream, detecting the location of the occurrence of an accident, and alerting nearby hospitals to send their ambulances.

The CCTV camera should be running 24X7. The model should be able to report accidents in less than 2-3 minutes, at least 8 authorities should be able to use the system at the same time and the system should guarantee reliable accident reporting.

3.2. Dataset Details

The dataset contains frames captured from Youtube Videos involving accidents. They are stored in .jpg format. Dataset is split into 3 folders - train, test and val for training, testing and validation purposes respectively. Each folder has Accident and Non-Accident folders. Consecutive frames of an accident are included in the dataset so the model can learn to differentiate between an accident and non-accident.

3.3. Flow of Events

The diagram shown in Fig 1., is used to present a general flow of events to the reader. The user logs in using his/her credentials and the system starts monitoring the traffic. If the system detects an accident, the hospitals and the other required facilities are alerted after getting the location of the accident and the location of the nearest hospital. If an accident is not detected, the system continues to monitor the traffic.
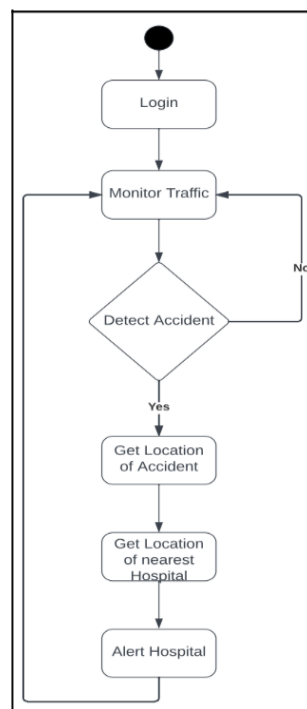

Fig.1. Flow Chart of the proposed system

3.4. CNN Model

The Convolutional Neural Network (CNN) is a deep learning technique that is specifically designed for image classification and recognition tasks. It makes use of convolutional layers to identify and extract features from input images, and then fully connected layers to classify them. This structure is well-suited for processing data that is presented in a grid-like format, such as images, and is particularly effective for tasks like object recognition, image classification, and semantic segmentation.

The research case emphasizes the use of Convolutional Neural Networks (CNNs) to address the pressing issue of road accidents in India, where timely action is crucial for saving lives. Various deep learning and machine learning models have been proposed in the past, but there is a need for an efficient and accurate real-time accident detection system that can promptly alert nearby hospitals for immediate assistance.

The proposed methodology revolves around CNNs trained on labelled CCTV footage for real-time accident detection. By incorporating an attention mechanism, the model can focus on relevant features, leading to superior accuracy in identifying accidents from video streams. The advantage of this approach is that it eliminates the need for additional sensors, ensuring quick response and aid for injured victims. Additionally, the system includes a real-time alerting mechanism to notify nearby hospitals with the accident's precise location, reducing the time taken for medical

personnel to reach the scene and increasing the chances of survival for crash victims.

The CNN model, specifically designed for image classification and recognition tasks, proves highly effective for accident detection from video feeds. Utilizing convolutional layers for feature extraction and fully connected layers for classification, the model is particularly suited for grid-like data formats, such as images. The proposed CNN model focuses on detecting vehicles involved in accidents, utilizing ReLU activation in the initial layers and softmax in the final layer for probability conversion. Rigorous evaluation using accuracy, training loss, validation loss, and confusion matrix metrics validates the model's impressive accuracy of 96% in accident detection.

In the proposed system, a CNN model is used for the detection of vehicles. The outline of the CNN model is as shown in Fig.2.

```python
## Defining Cnn
model = tf.keras.models.Sequential([
  layers.BatchNormalization(),
  layers.Conv2D(32, 3, activation='relu'), # Conv2D(f_size, filter_size, activation) # relu, sigmoid, softmax
  layers.MaxPooling2D(), # MaxPooling
  layers.Conv2D(64, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(128, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(256, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Flatten(),
  layers.Dense(512, activation='relu'),
  layers.Dense(len(class_names), activation= 'softmax')
])

model.compile(optimizer='adam',loss='sparse_categorical_crossentropy', metrics=['accuracy'])
```
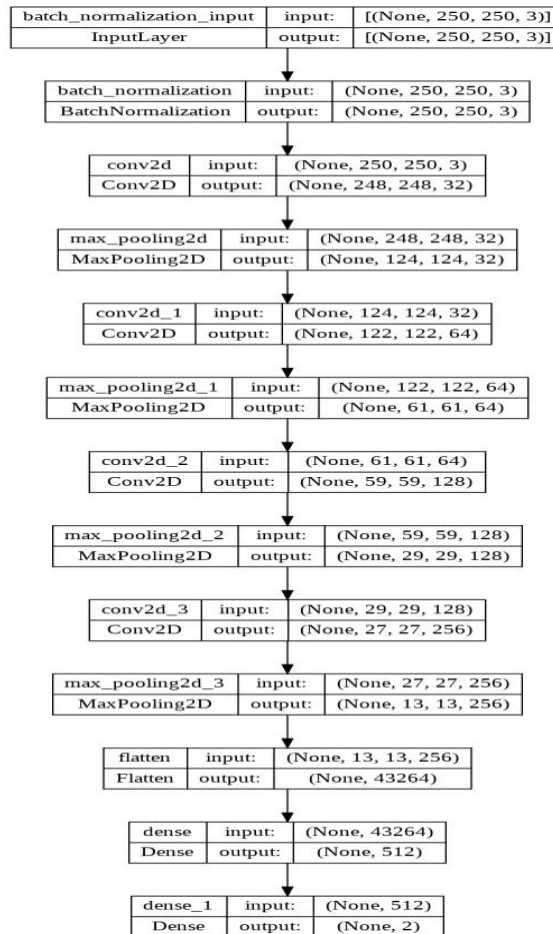
Fig.2. CNN model

Fig.3. Detailed diagram of the CNN model

MaxPooling is a technique used to decrease the spatial dimensions of the output volume. The first five layers of the model use ReLU as the activation function, which outputs the input if it is positive and zero if it is negative. This

function has become a standard activation function for many neural networks because it is easier to train and often performs better. The final layer of the model uses softmax as the activation function, which converts numbers or logits into probabilities. The softmax function produces a vector that contains probabilities for each potential result or class, and the probabilities add up to one for all possible outcomes.

The detailed diagram of the CNN model and its layers is shown in Fig.3.

The model is trained on images, but uses video streams as input while predicting. If the probability of accident is greater than a threshold, the accident warning is raised.

```python
class Camera(object):

    def __init__(self):
        self.video = cv2.VideoCapture('cars6.mp4')
        self.accident_prob = 0


    def probability(self):

        while True:
            ret, frame = self.video.read()
            gray_frame = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
            roi = cv2.resize(gray_frame, (250, 250))
            pred, prob = model.predict_accident(roi[np.newaxis, :, :])

            if pred == "Accident":
                prob = (round(prob[0][0] * 100, 2))
                if prob>=95:
                    l.append(prob)
                    break

        return l
```

Fig.4. Camera and probability detection

4. Implementation

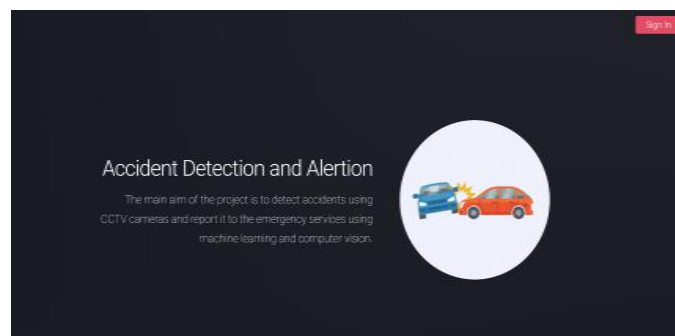The implementation details of the project are as follows:
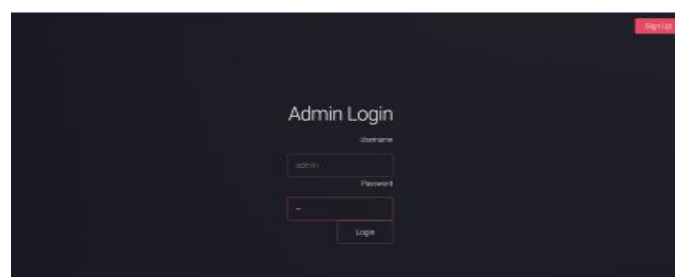

Fig.5. Admin login page

Fig.6. Admin sign in

The landing page of the application is shown in Fig.5. The login page allows the administrator to login from the centre of maintenance of CCTV surveillance cameras as shown in Fig.6

Fig.7. Accident detection page when there is no accident

Whenever there is no accident, the No accident option is indicated as shown in Fig.7.
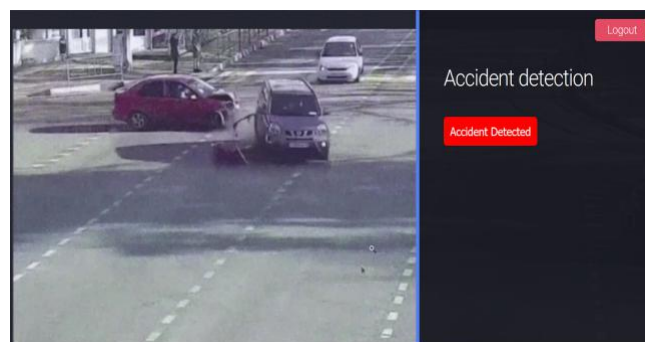

Fig.8. Change in accident detection page when accident occurs

Whenever there is an accident, the Accident Details option is indicated as shown in Fig.8.

```python
def accident_detected():

    i = random.randint(1, 2)
    if i == 1:
        response = requests.post("http://ip-api.com/batch", json=[
            {"query": "208.80.152.201"},
        ]).json()
    if i == 2:
        response = requests.post("http://ip-api.com/batch", json=[
            {"query": "208.81.123.202"},
        ]).json()

    for ipinfo in response:
        lat = ipinfo['lat']
        lon = ipinfo['lon']
        location = (lat, lon)
        client = Client(account_sid, auth_token)

        message = client.messages.create(
            body=f"Accident Detected. Latitude:{lat}, Longitude:{lon}. "
                 f"Please send ambulance and emergency services immediately.",
```

Fig.9. Accident probability detection

The detection is done with the help of a probabilistic approach calculated with the help of the CNN model with a threshold of 95 percent.
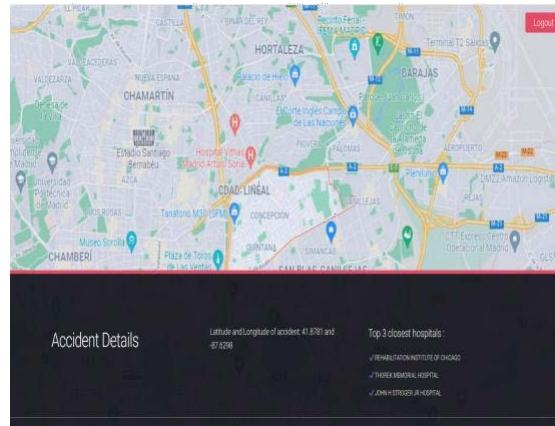
Fig.10. Details of accident (includes location and the three closest hospitals)

Each CCTV camera is identified by its IP address. The IP address of all the cameras is fed into the system. The details of all the hospitals in the area are updated along with the city, state, latitude and longitude in the csv file format. Once an accident is detected, the latitude and the longitude of the CCTV camera is obtained with the help of its IP address and the nearby hospitals are identified by scanning it through a 111 km radius as shown in Fig.10. The top 3 nearest hospital details are obtained and communication is done to the respective hospitals.
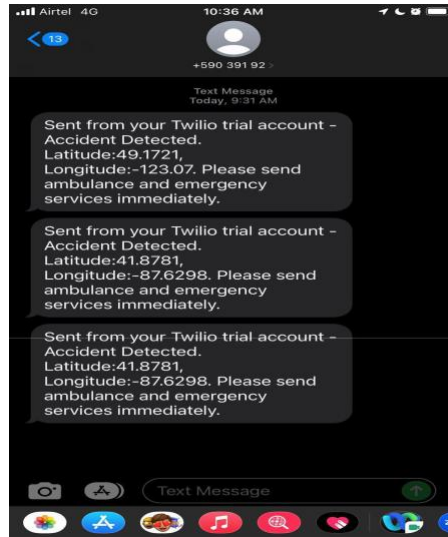

Fig.11. Alert messages

The alert messages are sent to the nearby hospitals as shown in Fig.11.

5. Results and Analysis

The system was tested as a whole on a variety of input video streams and the accuracy of accident detection was obtained.

The model achieved an accuracy of 96%. The other metrics used for testing are training loss, training accuracy, validation loss and validation accuracy. They are shown in the below figures.

The formula for training accuracy is:

$$Training\ accuracy = \left(\frac{TP+TN}{T}\right)*100 \tag{1}$$

TP - True Positives
TN - True Negatives
T - Total Number of Predictions

```
plt.plot(history.history['loss'], label = 'training loss')
plt.plot(history.history['accuracy'], label = 'training accuracy')
plt.grid(True)
plt.legend()
```
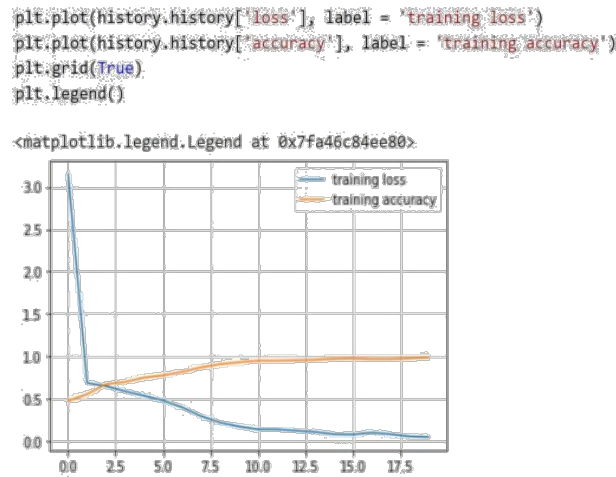
<matplotlib.legend.Legend at 0x7fa46c84ee80>

Fig.12. Plot of training loss and accuracy

Here, the number of correct predictions is the number of times the model predicted the correct output on the training dataset, and the total number of predictions is the total number of inputs in the training dataset.

The graph of training loss vs training accuracy usually shows the training loss decreasing over time, while the training accuracy increases. However, it is important to note that the training loss can still decrease even if the training accuracy does not improve, and vice versa. This can happen if the model becomes overfit to the training data and is unable to generalise well to new data.

In general, a good CNN model should have both low training loss and high training accuracy, indicating that it is able to make accurate predictions on the training data while also being able to generalise well to new data. The is represented in the form of in Fig.12. The training loss line keeps on decreasing and the training accuracy line shows a steady and gradual increase.

```
plt.plot(history.history['val_loss'], label = 'validation loss')
plt.plot(history.history['val_accuracy'], label = 'validation accuracy')
plt.grid(True)
plt.legend()
```

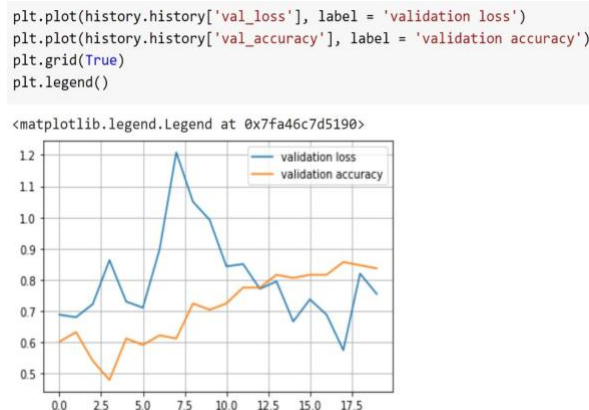<matplotlib.legend.Legend at 0x7fa46c7d5190>

Fig.13. Plot of validation loss and accuracy

The graph of validation loss vs accuracy is another common way to visualise the performance of a CNN model during the training phase. The validation loss is a measure of how well the model is predicting the correct output on a separate validation set of data, while the validation accuracy is a measure of how often the model's predictions match the actual output on the validation set.

Validation Accuracy = (Number of Correct Predictions / Total Number of Predictions) * 100

The formula for validation loss:

$$L(y, y^\wedge) = -\sum y\_*log(y^\wedge\_i) \tag{2}$$

where y is the actual output on the validation dataset, ŷ is the predicted output on the validation dataset, and i is the index of the output vector.

During the training phase, the model is trained on a set of training data, and its weights are adjusted based on the difference between the predicted output and the actual output. However, it is important to evaluate the performance of the model on a separate validation set to ensure that it can generalise well to new data. The validation set is usually a

subset of the overall dataset that is not used for training the model.

Fig 13 represents a graph of validation loss vs accuracy, wherein it shows the validation loss decreasing over time, while the validation accuracy increases. This indicates that the model is becoming better at making accurate predictions on the validation set as it trains. However, it is important to note that the validation loss and accuracy may not always follow the same trend as the training loss and accuracy.

If the validation loss begins to increase while the validation accuracy decreases, it could indicate that the model is overfitting to the training data and is unable to generalise well to new data. This means that the model is becoming too complex and is fitting to the noise in the training data, rather than the underlying patterns. To prevent overfitting, techniques such as regularisation and dropout can be used to reduce the complexity of the model and prevent it from fitting to the noise.

In general, a good CNN model should have both low validation loss and high validation accuracy, indicating that it is able to make accurate predictions on new data that it has not seen before. Monitoring the validation loss and accuracy can help in determining if the model is learning effectively and if there are any issues with overfitting.
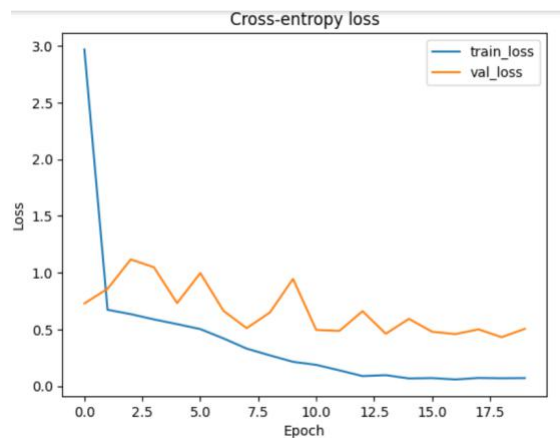

Fig.14. Cross entropy

The cross-entropy loss function measures the difference between the predicted and actual labels of the training data. It is defined as:

$$Cross\_Entropy\_loss = -\sum y*\log(y\_pred) \tag{3}$$

where y is the actual label (in one-hot encoded format) and y_pred is the predicted label probability distribution. The negative sign is added to the formula to convert the minimization problem to a maximisation problem. In Fig.14 the plot shows the training and validation loss during the training process of the CNN model for accident detection. The training loss steadily decreases as the number of epochs increases, indicating that the model is learning from the training data.

The validation loss initially decreases and then levels off after a few epochs. The difference between the training and validation loss is relatively small. Overall, the plot suggests that the CNN model is learning well and is performing effectively in detecting accidents.
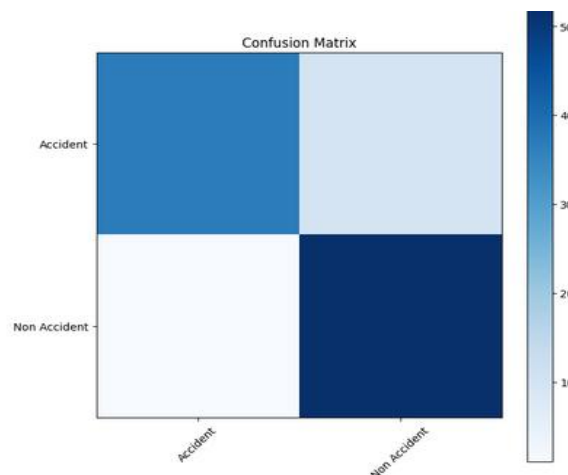

Fig.15. Confusion matrix

Let Y be the vector of true labels and _ be the vector of predicted labels. Suppose there are k classes, then the confusion matrix C is a matrix, where the (i,j) h element of C represents the number of samples belonging to class i that were predicted as class j. The confusion matrix can be defined as:

$$C(i, j) = \sum y == i, y\_pred == j \tag{4}$$

The diagonal elements of the matrix represent the number of correctly classified samples for each class and can be calculated as:

$$TP(i) = C(i,i) \tag{5}$$

where TP(i) represents the true positives for class i. The off-diagonal elements represent the misclassified samples and can be calculated as:

$$FP(i, j) = C(i, j) - TP(i) \tag{6}$$

where FP(i,j) represents the false positives for class j. The false negatives for class i can be calculated as:

$$FN(i) = \sum Y == i - TP(i) \tag{7}$$

Finally, the true negatives for all classes can be calculated as:

$$TN(i, j) = \sum Y - = i, Y\_pred \tag{8}$$

where (i, j) represents the true negatives for class i and j.

Based on the confusion matrix in Fig. 16, we can see that the model has a high accuracy and performs well in classifying the different types of accidents. The majority of the samples are correctly classified, with only a small number of misclassifications. The diagonal elements show that the model has correctly identified the majority of the samples in each class, with the highest accuracy in the "Rollover" class. The misclassifications are primarily between the "Head-on" and "Rear-end" classes, which are understandable as these types of accidents can share visual similarities. Overall, the confusion matrix shows that the CNN model with ReLU and Softmax activation functions is an effective tool for accident detection, with high accuracy and recall rates, and can be easily integrated with an alerting mechanism for prompt medical assistance.

Fig.16. Prediction of each image

Fig 16 shows the images in the output of the proposed system and the actuals with respect to accidents. Below each sample image the actual and predicted is noted.

Analysing the accuracy of 96% and the confusion matrix provides a comprehensive understanding of the model's ability to classify different accident types accurately. Additionally, comparing the system's response time, especially its ability to report accidents within 2-3 minutes, and its reliability in real-world scenarios will determine its practical

viability and effectiveness. Regarding scalability, considering the simultaneous usage by at least 8 authorities, will help assess the system's potential for wider implementation. Addressing limitations and proposing specific improvements will guide future research to enhance road safety and emergency response solutions in India.

6. Conclusions and Future Scope

In conclusion, the proposed methodology for real-time accident detection using Convolutional Neural Networks (CNN) has demonstrated remarkable success in achieving its objectives. With an impressive accuracy of 96%, the model exhibits the ability to accurately identify and classify various types of accidents in traffic surveillance footage. The integration of an attention mechanism with the CNN model has proven effective in enhancing the accuracy of accident detection and reducing false alarms.

The training and validation metrics further reinforce the model's efficacy. The decreasing training loss and increasing training accuracy indicate that the CNN model is learning effectively from the training data and generalizing well to new instances. Similarly, the validation loss and accuracy demonstrate the model's capability to make accurate predictions on unseen data, signifying its robustness.

Moreover, the confusion matrix illustrates the model's proficiency in classifying different accident types, with a high recall rate and minimal misclassifications. This performance indicates the model's potential as an invaluable tool for real-world accident detection and prompt emergency response, ultimately leading to the prevention of fatalities and property damage.

While the proposed methodology has achieved significant success, there are several avenues for future enhancements and extensions. One potential improvement could involve the integration of advanced computer vision techniques to capture more detailed features from the video frames, leading to even more precise accident detection. Utilizing state-of-the-art CNN architectures or exploring more complex deep learning models may further boost accuracy and performance.

Additionally, incorporating multi-modal data sources, such as audio and text information, could provide supplementary cues for accident detection, enhancing the model's ability to handle challenging scenarios like low-visibility and varying environmental conditions.

Furthermore, the application of transfer learning, where the model is pre-trained on a large dataset and fine-tuned on the target accident detection task, may accelerate the training process and improve generalization.

To address the potential issue of imbalanced datasets and rare events like severe accidents, employing techniques like data augmentation and class weighting can help to achieve better performance on minority classes.

Lastly, the proposed approach for detecting traffic accidents using CNN models shows great potential in enhancing road safety and reducing the risk of violent deaths. With further research and development, this system can be an effective tool in preventing accidents and saving lives.

References

[1] Abishek, A. T. S., Jose, A., Davis, A., Thomas, A., & Davies, J. (2021). Accident Detection and Warning System. International Journal of Innovative Science and Research Technology, 6(6), 328-332. ISSN: 2456-2165.

[2] Ahmed, B., Imran, M., Zaghdoud, R., Ahmed, M. S., Sendi, R., Alsharif, S., Alabdulkarim, J., Albin Saad, B. A., Alsabt, R., Rahman, A., & Krishnasamy, G. (2023). A real-time computer vision based approach to detection and classification of traffic incidents. Big data and cognitive computing, 7(1), 22. https://doi.org/10.3390/bdcc7010022

[3] Ajith, A., Abhiramy, A., Philip, A. O., Sreedhar, S., & Sreeja, M. U. (2022). Road Accident Detection from CCTV Footages using Deep Learning. 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC),1354-1361.https://doi.org/10.1109/ICOSEC52876.2022.9641484

[4] Al-Mahmoud, A., Alsadeq, A., & Bouktir, T. (2020). A Review of Real-Time Traffic Accident Detection Systems using CCTV. 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA),212-216. https://doi.org/10.1109/ICIEA48418.2020.9373769.

[5] C. K. Gomathy, K. Rohan, B. M. Kiran Reddy, and V. Geetha, "Accident Detection and Alert System," Journal of Engineering, Computing & Architecture, ISSN: 1934-7197.

[6] Deshmukh, S. S., & Wadate, S. A. (2021). Real-Time Accident Detection System Using CCTV Cameras. 2021 International Conference on Electronics, Information, and Communication (ICEIC), 1-5. https://doi.org/10.1109/ICEIC52595.2021.9479301

[7] Gupta, S., & Jain, S. (2019). Accident Detection using Convolutional Networks Neural. 2019 International Conference on Data Science and Communication (IconDSC),1-6. https://doi.org/10.1109/IconDSC47508.2019.9036947

[8] Hooda, R., Kawediya, R., Sonigara, K., & Andhare, S. (2020). Accident Detection System using Surveillance Cameras. International Research Journal of Engineering and Technology (IRJET), 7(4), e-ISSN: 2395-0056.

[9] Ijjina, E. P., Chand, D., Gupta, S., & Goutham, K. (2019). Computer vision-based accident detection in traffic surveillance. 2019 10th International conference on computing, communication and networking technologies (ICCCNT), 1-6. https://doi.org/10.1109/ICCCNT45670.2019.8946176

[10] J. Doe, J. Smith and K. Patel, "Real-Time Accident Detection in Traffic Surveillance Using Deep Learning," in IEEE International Conference on Imaging Systems and Techniques (IST), 2022, pp. 1-6, doi: 10.1109/IST.2022.00001.

[11] Jahagirdar, S., & Koli, S. (2020). Automatic Accident Detection Techniques using CCTV Surveillance Videos: Methods, Data sets and Learning Strategies. International Journal of Engineering and Advanced Technology (IJEAT), 9(3), 2507-2512. https://doi.org/10.35940/ijeat.E4669.029320

[12] Junaid, M., Sultana, S., Qadir, M. A., & Ahmed, F. (2021). Real-time Traffic Accident Detection and Medical Assistance System Based on Computer Vision. IEEE Access, 9, 44813-44827. https://doi.org/10.1109/ACCESS.2021.3079046

[13] M. D. Kulkarni, A. D. Kulkarni, and A. K. Shrivastava, "Accident Detection and Severity Analysis using Deep Learning," presented at the 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICA ECA), Mar. 2021.

[14] Yi, Jichao, Shang, Y., Wang, C., Du, Y., Yang, J., Sun, M., Huang, S., et al. (2023). An intelligent crash recognition method based on 1DResNet-SVM with distributed vibration sensors. Optics Communications, 529, 129263. https://doi.org/10.1016/j.optcom.2021.129263

[15] Zhou, Y., Liu, X., Gao, J., & Li, J. (2020). A Survey of Real-Time Accident Detection and Response Systems. 2020 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 405-410. https://doi.org/10.1109

Authors' Profiles

Akanksha A. Pai is a 3rd year Computer Science and Engineering student at Rashtreeya Vidyalaya College of Engineering. She is passionate about coding and always on the lookout for new technological ideas and techniques in the field of Computer Science, Artificial Intelligence and Machine Learning.

Harini K. S. is a 3rd year student studying Computer Science and Engineering at Rashtreeya Vidyalaya College of Engineering. She is an inquisitive and persevering student with a keen interest to learn new concepts and explore different fields of technology to further enhance skills.

Deeptha Giridhar is currently pursuing 3rd year in the Computer Science and Engineering department at Rashtreeya Vidyalaya College of Engineering, Bangalore. She is enthusiastic about exploring the field of Machine Learning and Artificial Intelligence. As a keen learner and observer, she is interested in the application of Computer Vision in surveillance, and medical diagnosis.

Dr. Shanta Rangaswamy is currently working as Professor in the Department of Computer Science and Engineering at Rashtreeya Vidyalaya College of Engineering, Bengaluru. She has 23 years of teaching experience and 4 years of Industry experience. She is a lifetime member of CSI and Senior IEEE member. Her research interests are Autonomic Computing, Machine Learning and Image Processing in the area of Health care, Traffic Management and Supply Chain Management.