TRAFFIC ACCIDENT DETECTION FROM SURVEILLANCE VIDEOS: AN AUTOMATED APPROACH FOR ENHANCED ROAD SAFETY

A MINI PROJECT REPORT

submitted by

EVELIN MANOJ

(TCR22CSCE07)

to

the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree

of

Master of Technology

in

Computer Science and Engineering

Department of Computer Science and Engineering

Government Engineering College

Thrissur

SEPTEMBER 2023

DECLARATION

I undersigned hereby declare that the mini project report **"Traffic Accident Detection From Surveillance Videos: An Automated Approach For Enhanced Road Safety"**, submitted for partial fulfillment of the requirements for the award of the degree of Master of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of **Prof. Rahmathulla K**, Assistant Professor Department of CSE. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place: Signature

Date: EVELIN MANOJ

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

GOVERNMENT ENGINEERING COLLEGE,THRISSUR

CERTIFICATE

This is to certify that the report entitled **”Traffic Accident Detection From Surveillance Videos: An Automated Approach For Enhanced Road Safety”**, submitted by **EVELIN MANOJ** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Computer Science and Engineering is a bonafide record of the seminar carried out by her under my guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Internal Supervisor(s) **Project Coordinator**

Head of the Department

ACKNOWLEDGEMENT

I give honour and praise to the lord who gave me wisdom and enabled me to complete this mini project successfully.

I would like to thank the Department of Computer Science and Engineering, Government Engineering College, Thrissur, for giving me the opportunity to present this mini project.

I would also like to thank **Dr. Shibily Joseph**, HOD, Department of Computer Science and Engineering for rendering her help and guidance during the completion of the mini project.

I would like to express my deep gratitude to our PG Co-ordinator **Dr.Salim A** for rendering his help and guidance throughout the mini project.

I would like to thank our mini project coordinator **Dr. Ajay James** for his sincere support throughout the mini project.

I would also like to thank my guide **Prof. Rahmathulla K** for his timely suggestions encouragement and guidance through the processes involved in the presentation of the mini project and the preparation of the report.

Last but not least, I express my heartfelt thanks to my friends and parents for their support and encouragement during this endeavor.

ABSTRACT

The frequency of traffic accidents is rising, which emphasizes the significance of quick and efficient accident detection systems to enhance traffic safety and emergency response. To save lives and reduce damage after an accident on one of our roads, quick action is necessary. Traditional accident detection techniques sometimes fail to send out signals in a timely manner. Imagine a future where security cameras do more than just monitor our surroundings; they also actively seek to keep us safe. By seamlessly integrating with security cameras, recording video streams, and extracting frames for in-the-moment accident categorization using a Convolutional Neural Network (CNN) architecture, this model realizes this objective.

Convolutional Neural Networks (CNNs), a subclass of deep neural networks popularised for their efficiency in image processing applications, are utilized by the proposed model. The main goal of this project is to create a CNN-based model capable of real-time video stream analysis and accident visual cue identification. This model collects frames from security camera footage and processes them using CNN architecture to perform accident categorization. It estimates the likelihood of an accident and makes an accident prediction. The system emits a buzzer sound when the likelihood is more than 90%.

Keywords: *Convolutional Neural Networks (CNNs), Real-time accident detection, Closed Circuit Television(CCTV), ReLu, softmax.*

CONTENTS

List of Figures

List of Abbreviations

AI	Artificial Intelligence
CCTV	Closed Circuit Television
CNN	Convolutional Neural Networks
DCNN	Deep Convolutional Neural Networks
OpenCV	Open Source Computer Vision Library
ReLu	Rectified Linear Unit

CHAPTER 1

INTRODUCTION

Traffic accidents, ranging from minor collisions to severe crashes, continue to be a pervasive and costly problem on a global scale. These incidents result in loss of life, injuries, property damage, and significant societal and economic burdens. Addressing this challenge requires not only preventive measures but also efficient response mechanisms. Traffic accidents, ranging from minor collisions to severe crashes, continue to be a pervasive and costly problem on a global scale. These incidents result in loss of life, injuries, property damage, and significant societal and economic burdens. Addressing this challenge requires not only preventive measures but also efficient response mechanisms.

Traffic accident detection is at the forefront of leveraging technology to enhance road safety. It seeks to develop systems and solutions that can rapidly identify and report accidents, reducing response times and ensuring that appropriate measures are taken promptly. The integration of cutting-edge technologies, such as computer vision, artificial intelligence, and sensor networks, has revolutionized the way we approach accident detection. Beyond the immediate response, traffic accident detection systems contribute to minimizing the overall impact of accidents. They aid in traffic management, rerouting vehicles away from accident-prone areas, and collecting valuable data for analysis. This data-driven approach supports efforts to prevent future accidents by identifying patterns and risk factors.

The economic and societal benefits of effective traffic accident detection are substantial. Reduced medical costs, lower insurance premiums, improved traffic flow, and decreased downtime all contribute to a more efficient and sustainable transportation system. Traffic accident detection is not a standalone solution but an integral part of a holistic approach to road safety. It complements preventive measures, such as traffic regulations and vehicle safety standards, by providing a responsive layer that addresses accidents as they occur. As we delve deeper into the world of traffic accident detection, we uncover a realm of innovation and potential. This exploration will delve into the technologies, methodologies, and real-world applications that are reshaping our approach to road safety. By embracing the power of technology, we aim to create roadways that are not only efficient but, above all, safe for everyone who travels them. The key contributions of the study are as follows:

1. A comprehensive system model to detect road accidents.
2. Predict their probability.
3. Investigation of AI and computer vision-based approaches for object and event detection.
4. Testing and validating the proposed model by contrasting it with the state-of-the-art techniques.

CHAPTER 2

LITERATURE SURVEY

In this section, some works that have been done in the same domain over the past decades are discussed.

The paper titled ”Vehicle detection and counting in high-resolution aerial images” is an integral piece of research that adds to the scientific knowledge base in the fields of vehicle surveillance, traffic management, and computer vision[1]. The authors focused on the advancement of vehicle detection technology, dealing with the challenges associated with previous methodologies, which could not promptly detect moving vehicles, especially in conditions of insufficient illumination. They developed their method by utilizing high-definition images from an original dataset they created featuring cars, buses, and trucks. These images covered a variety of scenes and lighting conditions, making them more representative of the real world. The study unveiled the successful application of Fast R-CNN for vehicle detection, albeit with a noted disadvantage of time consumption during the selection process of candidate frames. However, the work also highlights the persistent difficulties with detecting the type of vehicle, issues with edge extraction under low illumination, and the challenge of false detections when using aerial view angles. However, these concerns are balanced out by the method’s high precision and recall rates, indicating the technique’s productivity and advantageous contribution despite its limitations.

Xiaohui Huang et al. present an insightful approach to real-time traffic accident detection using a two-stream convolutional network[2]. The authors innovatively apply DeepSORT, with enhancements in tracking through the use of cosine metric learning. Their system's ability to handle both moving and static objects provides an edge over existing systems that can only track moving objects. The authors impressively demonstrate the execution of three subtasks - object detection, multiple object tracking (MOT), and near-accident detection - at a high frame rate, making it apt for real-time applications. The addition of their newly proposed Traffic Near-accident Detection (TNAD) dataset for evaluation is also a valuable contribution to the research community. The potential improvements could include rigorous testing under diverse real-world conditions such as varied lighting and adverse weather. Also, elaboration on how the model handles distortion introduced by the fisheye lens used, and clarification on the system's false-positive rates would add depth to the research.

Durgesh Kumar Yadav et al. in the study of "Accident detection using deep learning" provides an overview of three software designed for machine learning and computer vision applications: OpenCV, Keras, and TensorFlow[3]. OpenCV (Open Source Computer Vision Library) is an expansive open-source library enabling image processing, machine learning, and computer vision tasks. It contains a vast collection of over 2500 structured algorithms that cater to various tasks, from tracking camera movements to facial recognition. The versatility and generality of applications make OpenCV a valuable tool for both businesses and academic research. Keras, a neural network library written in Python, is designed to run on TensorFlow. It is designed to be fast, user-friendly, and modular, allowing for the convenient building and defining of models. Its goal is to simplify the process of building deep learning models by handling low-level APIs such as creating computational graphs, tensors, or other variables. TensorFlow wasn't discussed in detail in the provided text, however, it's known as a powerful open-source library for machine learning developed by Google Brain. Utilized by Keras as a backend, it excels in numerical operations and building large-scale neural networks. While providing useful insights, the paper doesn't delve into the detailed usage of these software tools. It describes their features and merits effectively, but additional information regarding their limitations, detailed functions, and a case study or practical application using this software would have been beneficial for a comprehensive understanding.

Da Yang et al. present a new model for traffic accident detection and classification, leveraging multi-vehicle trajectory data and a Deep Convolutional Neural Network (DCNN) model[4]. The DCNN model provides significant improvements in terms of accuracy, compared to existing models in the field. It achieves up to 100% accuracy for accident detection and up to 95% for accident classification. The model goes beyond the sole use of positional information of involved vehicles, capturing the dynamic changes in accident vehicles' trajectories over time. This allows not just for the prompt recognition of an accident, but also accurate identification of the accident type, with six types considered in this study. The performance of the model is assessed using data simulated with the PC-Crash software. This approach ensures that a broad spectrum of real-life accident scenarios is considered, enhancing the robustness of the results. Despite its strengths, the DCNN model's performance comparisons are only made against existing models rather than actual traffic accident data, and it is yet to be tested in real-world scenarios. Therefore, future research could benefit from testing this model in a real-world setting to assess its practical performance.

Okan Kopuklu et al. submitted a study on utilizing a deep contrastive learning approach for driver monitoring systems[5]. The authors have introduced the first multi-view, multi-modal Driver Anomaly Detection (DAD) dataset geared towards vision-based driver monitoring applications. One main highlight is the novelty of the deep contrastive learning approach, which effectively differentiates between normal and anomalous driving behaviors. Moreover, the authors thoroughly explore and elucidate performance parameters and workings of their approach through comprehensive ablation studies. The fusion of different views and modalities that improved performance is an intelligent step proving that various datasets contribute different yet complementary information for anomaly detection. Even though the DAD dataset might lack complete diversity of all possible anomalous actions and its dependence on multiple data sources could be a barrier in certain real-world applications, the study is nonetheless groundbreaking. It provides industry-important insights making it of significant relevance to those working on driver monitoring systems.

Chaeyoung Lee et al. introduced an interesting approach to improving traffic surveillance and management[6]. The proposed AI CCTV system aims to monitor and detect real-time traffic anomalies using deep learning models, marking a significant advancement over traditional CCTV systems. The study stands out for its innovative use of object detection models like YOLOv3 trained on real-world data. Leveraging AI allows for real-time detection of abnormalities on roads, potentially enabling swift responses by authorities and reducing traffic incidents. It does a commendable job of merging advanced technology with real-world applications. However, it's worth mentioning that the efficiency and effectiveness of the proposed system need to be further validated through robust testing and deployment in different traffic situations and settings. Despite this, the paper provides valuable insight into how AI can be harnessed for enhanced traffic surveillance and safety.

S. M. Sunny et al. explore an innovative approach to traffic management using image processing and analysis to detect, track, and classify vehicles within the context of a developing region such as Bangladesh[7]. The central proposition of a versatile Kit that not only identifies vehicles but also tracks their speed and identifies accident scenarios, is commendable. The integration of automated enforcement for traffic violations and reporting of emergencies promises a significant improvement in traffic regulation and safety. The paper is immaculately structured and clearly communicates the proposed system's functionality. This paper could be enhanced by addressing potential challenges including weather conditions, shadows, and large-scale implementation. There is a distinct lack of discussion on data security and privacy, considering the vast amounts of data captured, processed, and potentially stored by the proposed system.

Yu Tian et al. offer a thoughtful approach to video anomaly detection using Robust Temporal Feature Magnitude learning (RTFM)[8]. The proposed model is theoretically robust, inspired by top-k instance multiple instance learning (MIL), and adept at handling both short and long-range temporal dependencies with the use of Pyramid Dilated Convolutions (PDC) and Temporal Self-Attention (TSA) mechanisms. What sets RTFM apart is its unique focus on feature magnitude for discerning abnormalities rather than conventional separability presumptions, making it a novel development in the field. The method displays impressive performance, securing state-of-the-art results on several benchmark datasets. Despite this, the paper could have benefited from an in-depth comparative study exploring the specifics of RTFM's strengths and weaknesses in contrast to other contemporary models. Also, more clarity on its computational complexity metrics, ability to handle extremely large datasets or real-time anomaly detection utility would add value.

Thakare Kamalakar Vijay et al. focus on implementing a model for video-based detection of road accidents using deep learning techniques[9]. The researchers extract features using C3D and I3D networks and implement the ViSiL network within their model. They experimented with segment sizes of 8, 10, and 16 frames but found no significant impact on performance. A 3-layer MLP classifier model, similar to the one proposed by Feng et al., is used with specific hyperparameters. The model is trained using a learning rate of 0.01 for 350 iterations with an Adagrad optimizer. The paper introduces a novel synthetic dataset, named MP-RAD, produced using a video game (GTA-V). The authors pursued cross-validation, revealing that the synthetically trained model outperformed baseline models when tested on real-world videos, and demonstrating the utility of synthetic samples for training purposes.

Kamalakar Vijay Thakare et al. propose a new architecture that demonstrates robust performance in video analysis, specifically in the context of road accident videos[10]. This is reflected in the high recall values achieved: 0.87 (with 2-body proposals) and 0.91 (with 3-body proposals). The proposed method shows superior performance to weakly-supervised methods. It demonstrates significantly larger margins in the AUC results on the CADP dataset, with gains ranging from 3% to 13%. This architecture also achieves a lower False Alarm Rate (FAR) in comparison to existing methods, making it potentially more reliable. The authors also employ a novel technique of cropping initial parts of accident videos to represent normal behavior, to overcome the challenge of the CADP dataset not having normal videos. This innovation allows them to train the weakly supervised methods.

Yifan Sui et al. enable automatic detection and analysis of traffic accidents using surveillance video, eliminating human intervention in the process and increasing efficiency[11]. The introduced system facilitates reliable estimations of speed and collision angles of vehicles involved in an accident by providing a vertical view of the trajectory. he system has been tested and worked well on both a computer and HiKey970, proving its versatility. This paper also highlighted a few potential limitations, such as decreased identification accuracy if a vehicle is blocked, missed accidents due to low-quality videos, and the non-utilization of HiKey970's NPU for acceleration due to the unavailability of the required API at the time.

Yajun Xu et al. introduced a novel dataset named Traffic Accidents Dataset (TAD), aimed at accident detection in the industry, particularly concentrating on highway scene situations[12]. This dataset was reported to be the largest-scale open-sourced data of its kind, incorporating four types of accidents across various real-life instances. In addition to providing the largest dataset on this subject thus far, the authors also subjected it to a thorough experimental evaluation by utilizing three mainstream visual computing algorithms. The methodology was selected to perform image classification, video classification, and object detection. Experimental evaluations demonstrated that TAD can act as a reliable benchmark for algorithms assessing different types of traffic incidents, proving its potential to stimulate further research and application in traffic accident detection within intelligent transportation systems.

Yihang Zhang and Yunsick Sung presented an experimental study on the performance of various traffic-accident-detection models, including DNN, LSTMDTR, and ViT-TA, and a proposed framework[13]. The training results are depicted through loss and accuracy convergence plots. Results show that the proposed framework shows better performance than the other models. Training took 1000 epochs, with the new framework's training loss starting at 0.65 and decreasing to 0.02, and accuracy increasing from 0.68 to 0.99. The other models saw slower convergence rates and lower accuracy, with DNN reaching a final loss of 0.29 and accuracy of 0.88, LSTMDTR reaching a loss of 0.46 and accuracy of 0.86, and ViT-TA reaching a loss of 0.46 and accuracy of 0.78. The proposed framework's superior performance demonstrates its effectiveness in detecting traffic accidents.

These authors also contributed an academic paper that delves into the wide-ranging applications of Convolutional Neural Networks (CNNs), from natural language processing tasks to cyber security. CNNs are powerful models capable of extracting and analyzing features from data, with their capabilities demonstrated by their successful use in sentiment analysis, machine translation, music classification, and emotion detection in music[14]. This paper presents the experimental results of a proposed method involving a computational model. The model's performance is evaluated in terms of training loss, validation loss, and accuracy. Initially, the training loss begins at approximately 2.88 and then slowly converges to around 0.06 after four epochs. The validation loss shows a similar trend starting at around 3.41 and converging to 0.12. The accuracy of the model also shows significant improvement as the epochs progress. The initial training accuracy is about 55.63, which eventually rises to about 97.13. Similarly, the validation accuracy starts around 77.87 and increases to 95.87 after further training. The study provides evidence on how the proposed method can successfully extract relevant features from influence maps to facilitate traffic accident detection.

Mohammed Imran Basheer Ahmed et al. proposed an AI and computer vision-based system model to aid in road safety[15]. It provides a comprehensive solution for detecting road accidents, their severity, and potential post-accident fires. The proposed system uses AI and computer vision-based approaches for object and event detection, presumably to recognize accidents and relevant events accurately. The model has been tested and validated against state-of-the-art techniques, ensuring its usability and effectiveness. The system has an alert mechanism and can initiate a post-accident emergency response, highlighting its practical value in real-world scenarios. The exact technical aspects of the model, as well as how it was tested and validated, are further discussed in the main body of the paper, which is not included in the given excerpt.

CHAPTER 3

THE PROPOSED SYSTEM

Road safety has been degraded as a result of the lack of a reliable, automated, and scalable system for accurate traffic accident identification from surveillance videos. Existing techniques frequently have inefficiencies, erratic results, and restricted scalability. An automated solution is essential because it has the ability to significantly cut reaction times, hence lowering the danger of future injury and accidents. It also promises to produce better accident data, enabling deeper analysis and proactive preventative measures. Convolutional Neural Networks (CNNs) have been used extensively in the field of image analysis and are well-known for their power. Deep learning methods, in particular CNNs, outperform traditional machine learning methods in terms of predicted accuracy. The proposed system relies on CNNs to revolutionize traffic accident prediction and, as a result, improve road safety in light of these developments.

3.1 The proposed Model

We explore the intricate details of the suggested traffic accident detection methodology and its underlying architecture in this part. The first step of the procedure is the capture of video data from CCTV sources, which is followed by the frame extraction stage, which is effectively carried out using Python's OpenCV module. Before entering the Convolutional Neural Network (CNN) model for accident prediction, each frame is transformed, shrinking to a uniform 250x250 pixel dimension. CNN architecture's final layer, which computes accident probability, uses the softmax activation function. Then, these probabilities are properly transformed into straightforward percentages. The system replies with a beep sound when the prediction probability exceeds a key threshold of 90%, instantly warning of the increased danger of an accident.

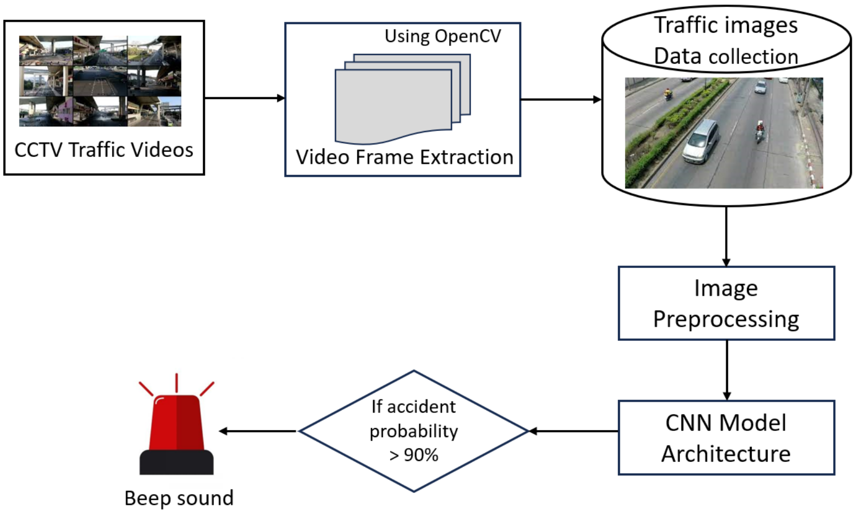

Figure 3.1: Proposed Model Architecture

3.1.1 Proposed CNN Model

The proposed system uses a CNN model which consists of four 2D convolutional layers of filter sizes 32,64,128 and 256 respectively and uses the ReLU (Rectified Linear Unit) activation function. Convolutional layers are responsible for learning spatial patterns in the data. It also adds four max-pooling layers that alternate with the convolutional layers. Max-pooling reduces the spatial dimensions of the output from the previous layer while retaining the most important information. The flatten layer flattens the output from the previous layers into a one-dimensional vector. This is necessary before passing the data to fully connected layers. The fully connected (dense) layer with 512 units and ReLU activation is responsible for learning high-level features and making predictions. The final layer uses the softmax activation function to compute class probabilities, making it suitable for multi-class classification tasks.

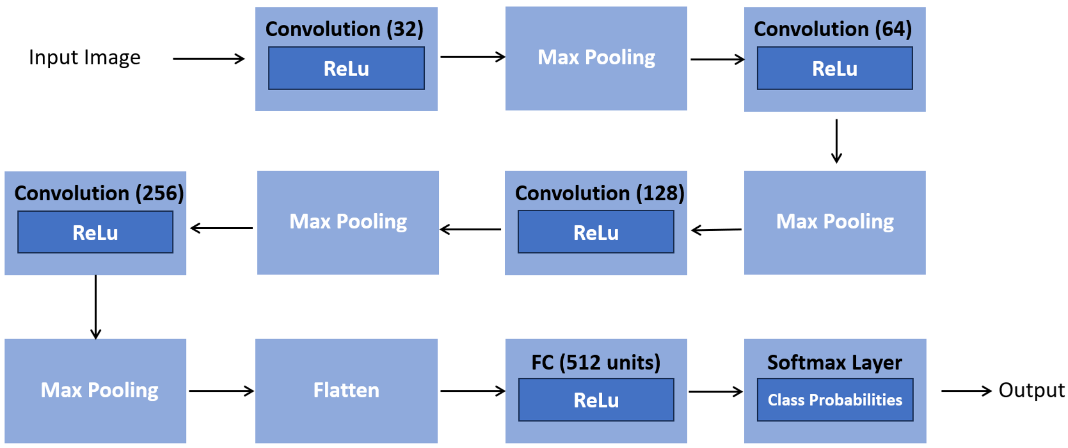

Figure 3.2: Proposed CNN Architecture

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 batch_normalization_1 (Batc  (None, 250, 250, 3)      12
 hNormalization)

 conv2d_4 (Conv2D)           (None, 248, 248, 32)      896

 max_pooling2d_4 (MaxPooling  (None, 124, 124, 32)     0
 2D)

 conv2d_5 (Conv2D)           (None, 122, 122, 64)      18496

 max_pooling2d_5 (MaxPooling  (None, 61, 61, 64)       0
 2D)

 conv2d_6 (Conv2D)           (None, 59, 59, 128)       73856

 max_pooling2d_6 (MaxPooling  (None, 29, 29, 128)      0
 2D)

 conv2d_7 (Conv2D)           (None, 27, 27, 256)       295168

 max_pooling2d_7 (MaxPooling  (None, 13, 13, 256)      0
...
Total params: 22,541,134
Trainable params: 22,541,128
Non-trainable params: 6
_________________________________________________________________
```

Figure 3.3: CNN Model Summary.

3.2 Materials and Method

This section will briefly explain the materials used and the working of different approaches used in this project.

3.2.1 OpenCV

The abbreviation OpenCV stands for ”Open Source Computer Vision Library,” and it is a robust and flexible open-source computer vision and image processing framework. OpenCV is a crucial tool when it comes to video frame extraction. It is an essential tool for many applications, including surveillance, object identification, and most significantly, traffic accident detection systems, since it enables the seamless recording and manipulation of video streams.

3.2.2 Dataset

”Accident Detection From CCTV Footage” dataset is used. It contains CCTV footage frames of accidents and non-accidents. The dataset is split into 3 folders - train, test and val. 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.

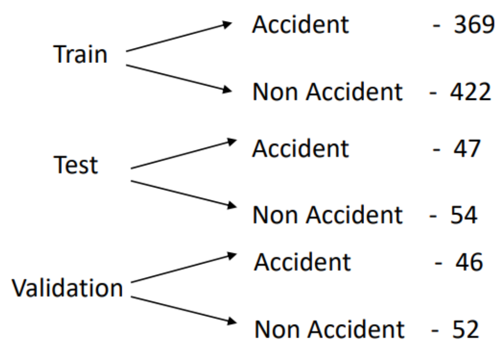

Figure 3.4: Dataset

3.2.3 CNN

CNNs are deep-learning neural networks that are mostly employed for feature extraction. It is done by applying filters to batches of data points. Feature extraction is done to reduce the number of features in a dataset by creating new features from existing ones. This new feature set preserves the information in the original dataset. A convolutional neural network has an input layer followed by a convolutional layer, pooling layers, fully connected layer, and an output layer. Figure 3.4 shows CNN architecture.

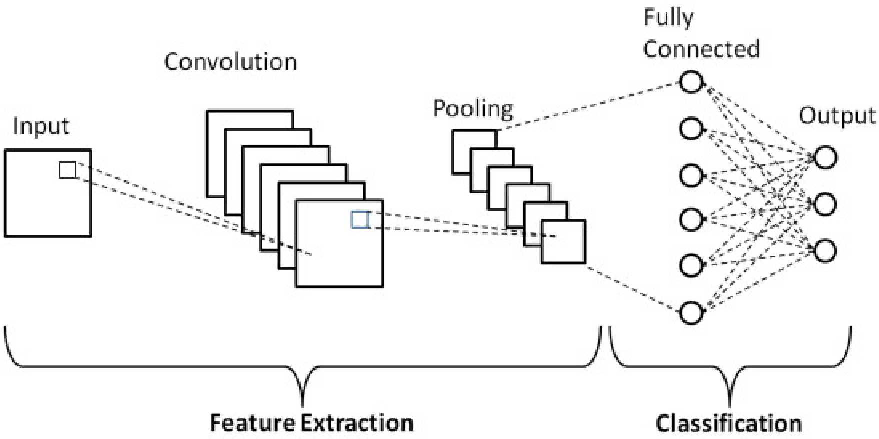

Figure 3.5: Basic CNN Architecture

3.2.3.1 Convolutional Layer

This layer is used to extract the various features from the input images. The mathematical operation of convolution is performed between the input matrix and a filter of a particular size MxN and generates a feature map. The following layer receives this output. The activation function is used to confine the feature map to a specific range. Convolutional operation is shown in the figure 3.5.

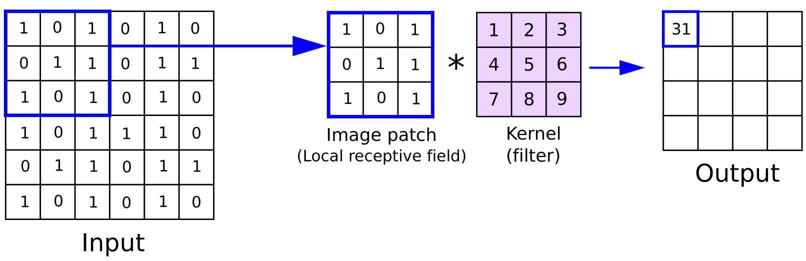

Figure 3.6: Convolutional Operation

3.2.3.2 ReLu Activation Function

ReLU operates by setting all negative input values to zero and leaving positive values unchanged. This means that if the input is greater than zero, ReLU outputs the input itself; otherwise, it outputs zero. This simple, piecewise-linear function introduces non-linearity to neural networks, enabling them to learn complex patterns and representations in data. ReLU is computationally efficient, helping networks train faster, and it mitigates the vanishing gradient problem, which can occur with other activation functions. These qualities have made ReLU a popular choice in modern neural network architectures, contributing to their success in various machine learning tasks.

3.2.3.3 Pooling Layer

The main aim of this layer is to downsample the feature map to summarize the features. This helps in decreasing the dimension of the feature map to reduce the computational costs. Max Pooling, is taken from the feature map. Average Pooling calculates the average of the elements in a predefined-sized matrix section. Pooling operation also helps to reduce the overfitting problem. An example of a pooling operation is depicted in Figure 3.6.

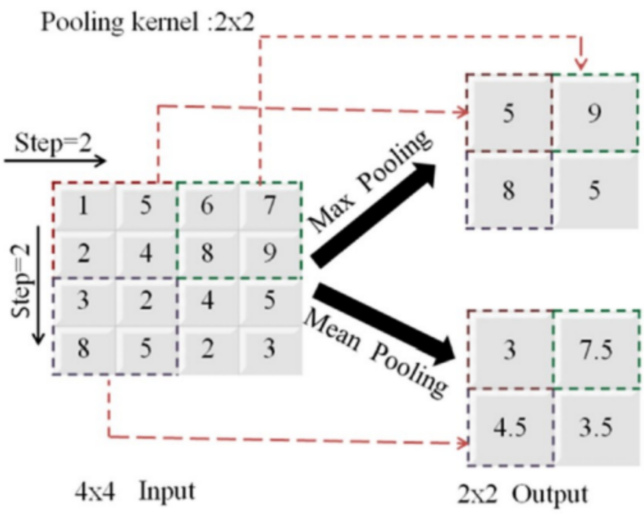

Figure 3.7: Modes of the Max Pooling Operation

3.2.3.4 Softmax

The softmax activation function takes a vector of real numbers as input and calculates the exponentials of each element in the vector. It then normalizes these exponentials by dividing each of them by the sum of all the exponentials. The result is a new vector of values between 0 and 1, where each value represents the probability that the input belongs to a particular class. In essence, softmax assigns higher probabilities to larger input values, helping the neural network make informed decisions about which class an input sample most likely belongs to. It is a fundamental component in the final layer of a neural network for multi-class classification tasks, ensuring that the class probabilities sum to 1, making it a suitable choice for predicting one of several mutually exclusive outcomes.

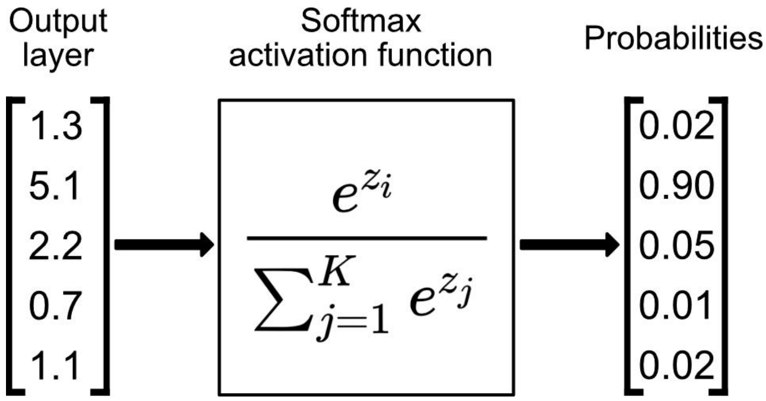

Figure 3.8: Softmax Activation Function

CHAPTER 4

EXPERIMENT AND RESULTS

The set of experiments was carried out using the Visual Studio Code platform. The proposed CNN model uses the Accident Detection From CCTV Footage dataset from Kaggle which contains CCTV footage frames of accidents and non-accidents. CNN is trained using 369 Accident images and 422 Non Accident images. The model was tested using the test folder and val folder which contain 47 Accident images - 54 Non Accident images and 46 Accident images - 52 Non Accident images respectively.

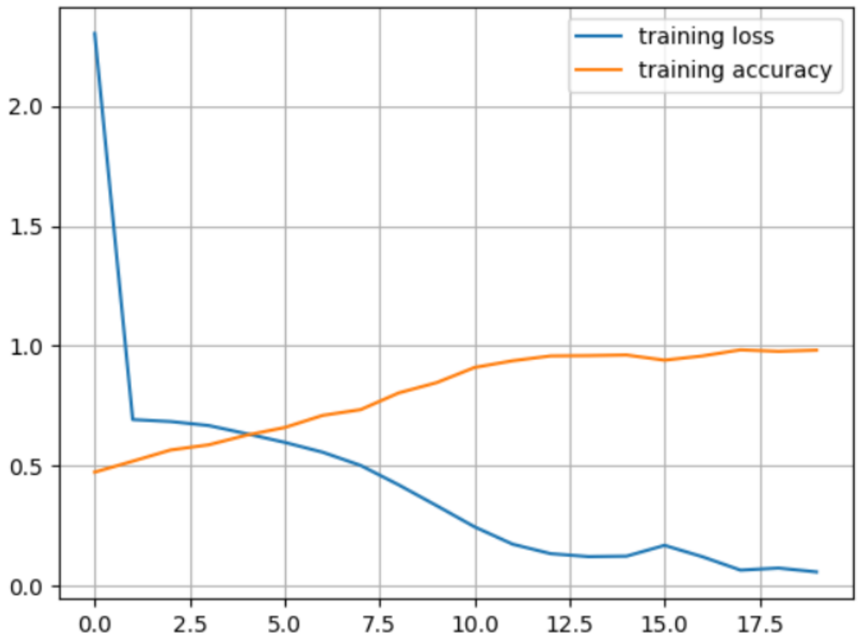

Figure 4.1: Plot for training loss and accuracy.

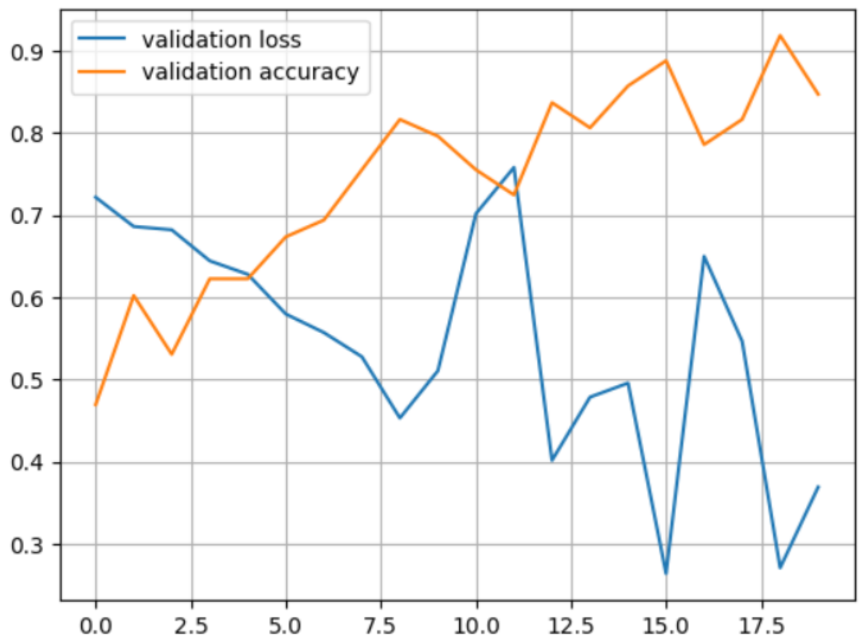

Figure 4.2: Plot for validation loss and accuracy.

Figure 4.3: Visualization of Test Results.

For real-time validation, the system's webcam and YouTube videos are used. Whenever the accident probability exceeded 90%, the beep sound was produced. The model's performance was evaluated based on the accuracy and loss values. After training 0.0582 loss and 0.9823 accuracy were obtained. As the validation result, the validation loss was 0.3691 and the validation accuracy was 0.8469.

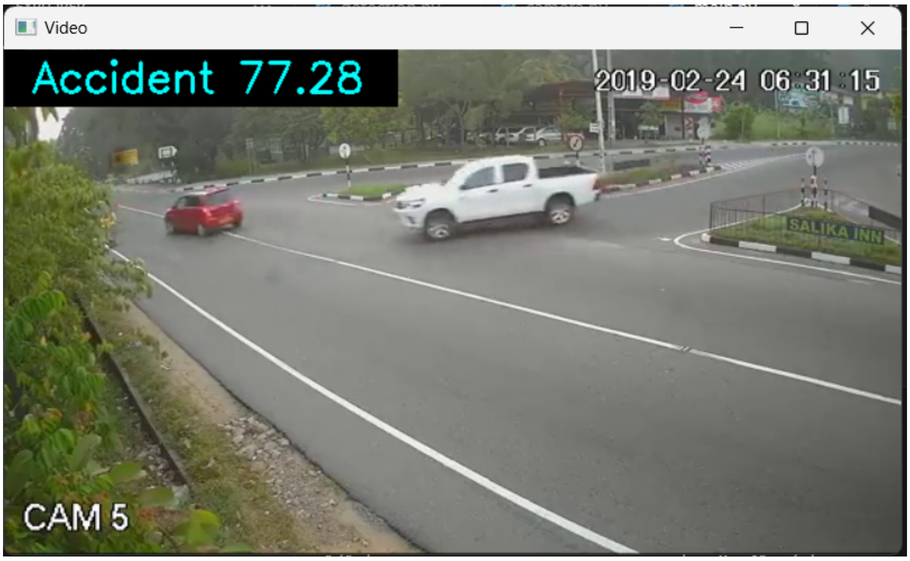

Figure 4.4: Detecting Non Accident.

Figure 4.5: Detecting Accident.

CHAPTER 5

CONCLUSION

This project is an example of innovation and development in the dynamic field of road safety. The mission was clear from the start: to use technology, in particular Convolutional Neural Networks (CNNs), to transform accident detection on our roads. This CNN-based model demonstrated impressive accuracy in accident detection, beating conventional approaches by a wide margin. It quickly and correctly foresaw probable collisions and provided important information about upcoming traffic risks. The real-time alertness of the suggested system is its distinguishing feature. It transformed the idea of surveillance cameras from passive observers to proactive defenders of traffic safety by seamlessly analysing video frames and generating quick predictions. It is possible that this real-time capacity may reshape emergency response times and eventually save lives. It has been demonstrated that adding a warning system that is activated when the accident risk approaches 90% is a game-changer. This provides the path for prompt involvement and help in addition to assisting in the initial accident recognition process.

Roadway safety improvement efforts are continuing. Future improvements, dataset extension, and model fine-tuning are in the works. In conclusion, the junction of technology and safety is at a turning point thanks to my real-time accident detection initiative. It's evidence of the impressive advancement I can make when I apply innovation to significant real-world problems. It is my hope that this technology will spur more developments in the field in addition to improving road safety.

.

REFERENCES

[1] Huansheng Song, Haoxiang Liang, Huaiyu Li, Zhe Dai, and Xu Yun. Vision-based vehicle detection and counting system using deep learning in highway scenes. *European Transport Research Review*, 11(1):51, Dec 2019.

[2] Xiaohui Huang, Pan He, Anand Rangarajan, and Sanjay Ranka. Intelligent intersection: Two-stream convolutional networks for real-time near-accident detection in traffic video. *ACM Trans. Spatial Algorithms Syst.*, 6(2), jan 2020.

[3] Durgesh Kumar Yadav, Renu, Ankita, and Iftisham Anjum. Accident detection using deep learning. In *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pages 232–235, 2020.

[4] Da Yang, Yuezhu Wu, Feng Sun, Jing Chen, Donghai Zhai, and Chuanyun Fu. Freeway accident detection and classification based on the multi-vehicle trajectory data and deep learning model. *Transportation Research Part C: Emerging Technologies*, 130:103303, 09 2021.

[5] Okan Kopuklu, Jiapeng Zheng, Hang Xu, and Gerhard Rigoll. Driver anomaly detection: A dataset and contrastive learning approach. 09 2020.

[6] Chaeyoung Lee, Hyomin Kim, Sejong Oh, and Illchul Doo. A study on building a "real-time vehicle accident and road obstacle notification model" using ai cctv. *Applied Sciences*, 11(17), 2021.

[7] S. M. Sunny, Towhidur Rahman, S. M. Zohurul Islam, Asif Mujtaba, Kazi Firoz Ahmed, and Shuvra Saha. Image based automatic traffic surveillance system through number-plate identification and accident detection. In

2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pages 467–472, 2021.

[8] Yu Tian, Guansong Pang, Yuanhong Chen, Rajvinder Singh, Johan W. Verjans, and Gustavo Carneiro. Weakly-supervised Video Anomaly Detection with Robust Temporal Feature Magnitude Learning. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4955–4966, Montreal, QC, Canada, October 2021. IEEE.

[9] Thakare Kamalakar Vijay, Debi Prosad Dogra, Heeseung Choi, Gipyo Nam, and Ig-Jae Kim. Detection of road accidents using synthetically generated multi-perspective accident videos. *IEEE Trans. Intell. Transp. Syst.*, pages 1–10, 2022.

[10] Kamalakar Vijay Thakare, Debi Prosad Dogra, Heeseung Choi, Haksub Kim, and Ig-Jae Kim. Object Interaction-Based Localization and Description of Road Accident Events Using Deep Learning. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):20601–20613, November 2022.

[11] Yifan Sui, Shaodong Zhou, Zhiyang Ju, and Hui Zhang. A Vision-Based System Design and Implementation for Accident Detection and Analysis via Traffic Surveillance Video. *IEEE Canadian Journal of Electrical and Computer Engineering*, 45(2):171–181, 2022.

[12] Yajun Xu, Chuwen Huang, Yibing Nan, and Shiguo Lian. TAD: A Large-Scale Benchmark for Traffic Accidents Detection from Video Surveillance, September 2022. arXiv:2209.12386 [cs].

[13] Yihang Zhang and Yunsick Sung. Traffic Accident Detection Using Background Subtraction and CNN Encoder–Transformer Decoder in Video Frames. *Mathematics*, 11(13):2884, June 2023.

[14] Yihang Zhang and Yunsick Sung. Traffic Accident Detection Method Using Trajectory Tracking and Influence Maps. *Mathematics*, 11(7):1743, April 2023.

[15] Mohammed Imran Basheer Ahmed, Rim Zaghdoud, Mohammed Salih Ahmed, Razan Sendi, Sarah Alsharif, Jomana Alabdulkarim, Bashayr Adnan Albin Saad, Reema Alsabt, Atta Rahman, and Gomathi Krishnasamy. A real-time computer vision based approach to detection and classification of traffic incidents. *Big Data and Cognitive Computing*, 7(1), 2023.