Volume 4, Issue II, February, 2025, No. 57, pp. 736-743
Submitted 2/2/2025; Final peer review 30/2/2025
Online Publication 3/3/2025
Available Online at http://www.ijortacs.com

DEVELOPMENT OF A DEEP LEARNING MODEL FOR THE PREVENTION AND CONTROL OF ROAD ACCIDENT

By
Engr. Dr. Harbor M.C.(Ph.D.)
Nigerian Defence Academy
Corresponding Author's Email and Tel: dumbormart@gmail.com , [1]+2348068215797

Abstract

This paper describes the development of a convolutional neural network (CNN) based vehicle accident prevention and control system which is developed to improve vehicle safety by providing real time predictions and control decisions. The model has been developed with several convolutional layers followed by Rectified Linear Unit (ReLU) activation and drop out to enhance the accuracy while avoiding overfitting. The CNN model was trained on the Honda Deep Drive (HDD) dataset and other locally collected tricycle data for the training of the model, 21,300 samples. The system has been developed and implemented in MATLAB's Simulink environment and performed effectively, with a detection rate of 98.1% and an ROC score of 0.98. It was found that the proposed system is better than the current state of the art models in terms of accuracy and reliability. These results have shown that the proposed CNN based system is efficient in accident detection and prevention when compared with the conventional methods. Due to its robustness and low computational complexity, it can be easily adopted for use in mobile and hardware environments for real time applications. Thus, this research is significant in the improvement of intelligent vehicle safety systems and provides a basis for future work on deep learning for autonomous and semi-autonomous vehicles.

Keywords: Accident Prevention; Convolutional Neural Network (CNN); Vehicle Safety System; Deep Learning; Feature Extraction

1. INTRODUCTION

According to the (National Highway Traffic Safet, 2023) research on transportation, over 900 million individuals travel by car all over the world every day. However, despite the importance of this great invention of mankind, it equally means that over 900 million people stand the risk of car or road accident daily. The evidence is there for all to see in the news, social media, along the express highways among others. According to the Road Traffic Accident (RTA) system over 1.2 million deaths and 50 million people are injured every year due to road crash (WHO, 2023; Aamir et al., 2019).

In recent years, the number of road accident increase has raised great concerns across the globe and presents a major challenge for road traffic administrators. If serious solution is not proposed, by the year 2025 accident will be the third major cause of human mortality rate, putting road safety ahead of HIV/AIDS, malaria, and other acts of violence (Benos et al., 2021). Various methods have been identified as the solution to this canker, which includes driving lesson schools and licensing, road safety rules and regulations, implementation of automatic braking systems, cruise control system, among others. However, despite the success, the rate at which accident happens keep increasing even more.

This problem has recently gain research attention with various techniques proposed already to solve the problem. Wu and Wang (2017) proposed the use of intelligent vehicle detection and tracking system, but the success is limited by the data structure which did not consider all classes of vehicles like the tri cycles machines for instance. Chen et al. (2017) presented vehicle detection system using building regional covariance descriptors. The limitation of the result is accuracy of 92% which can be increased. Kim et al. (2015) used histogram of oriented gradient for on road vehicle detection, however the technique lack common sense of the problem in question and have to be improved with artificial intelligence.

Other approaches use polynomial expansion based on motion estimation technique, local scale invariant feature technique, speed up robust feature, feature matching and Kalman filter technique among others (Mantripragada et al., 2014; Bay et al., 2016). However, despite the success, these techniques are not reliable enough for autonomous accident prevention system. Nevertheless, artificial intelligence (A.I) technique has been singled out to produce better result compared to other approaches.

Artificial intelligence is machines which mimic human behaviour. This are achieved using Machine Learning (ML) technique. The ML are series of mathematical algorithms used to make decisions based on the training dataset used. Today there are various ML algorithms, ranging from support vector machine, K-Nearest Neighbour, clustering technique, Convolutional Neural Network (CNN) among others. All have their advantages and disadvantages; however, the CNN performs best for image classification problem like the case study and is therefore proposed in this research to be used for training the problem under study and then make correct and intelligent decision. This system will then be integrated in autonomous vehicles as an accident prevention system, with high level of efficiency and system reliability.

1. DESIGN METHODOLOGY

In this research paper, a Convolutional Neural Network (CNN) model that features multiple output channels is proposed. The study first designed a model solely for predicting the speed of the ego-vehicle, and then extended its capability to provide predictions for not only the vehicle speed, but also the brake-pedal pressure and emergency-brake signals. The foundational architecture of the initial model is based on the utilization of pairs of images in conjunction with the corresponding ego-vehicle speeds, where each input image is associated with a specific distance value, expressed in kilometres per hour (km/h). Specifically, the input dataset consists of a sequence of 20 consecutive frames, each comprising images and their corresponding distance values. The CNN's primary function in this initial configuration is to estimate the current vehicle speed for the given frame.

1.1 Data Acquisition

There are many publicly available datasets for the development of vision-based self-driving vehicles, but few of them feature continuous image sequences (or videos) along with additional measurement data (such as heading, vehicle speed, brake sensor data, etc.).This is the main reason why this study used the Honda Deep Drive (HDD) (He et al., 2017) dataset that is available for research purposes. Also, data of tricycles were collected from the federal ministry of transportation. The sample size of data collected is 21,300 samples of tricycles and was stored in the system image repository to create the training dataset. The samples of the data collected were presented in the Figure 1;

Figure 1: Samples of the collected data

1.2 Feature Extraction

Having collected the data in image format, it was converted into statistical features using the Histogram of Feature Gradient (Sadek et al., 2010). The reason this feature extraction technique was adopted was due to its ability to correctly extract the rich features in an image and then convert into statistical equivalent for training purpose (Saxe and Berlin, 2015). While extracting the pairs of images with corresponding distance and brake-pedal pressure values, we opted to further examine the relation between the ego-vehicle distance, acceleration, and brake signals in order to identify and better understand potentially hazardous situations. We decided to identify hard-brake situations by analysing the dataset, and to generate an additional emergency-brake signal, which holds practical significance in potentially dangerous traffic scenarios. This signal is produced a few frames prior to the detection of an actual hard-braking event.

1.3 Convolutional Neural Network (CNN) Model

Based on a series of consecutive images, the original CNN model can accurately forecast a vehicle's distance (Itu and Danescu, 2024). Convolutional layers precede a fully connected layer in the CNN architecture. As a result, the model can directly predict the distance value from raw pixel data. In Figure 2, the model's structure is displayed.

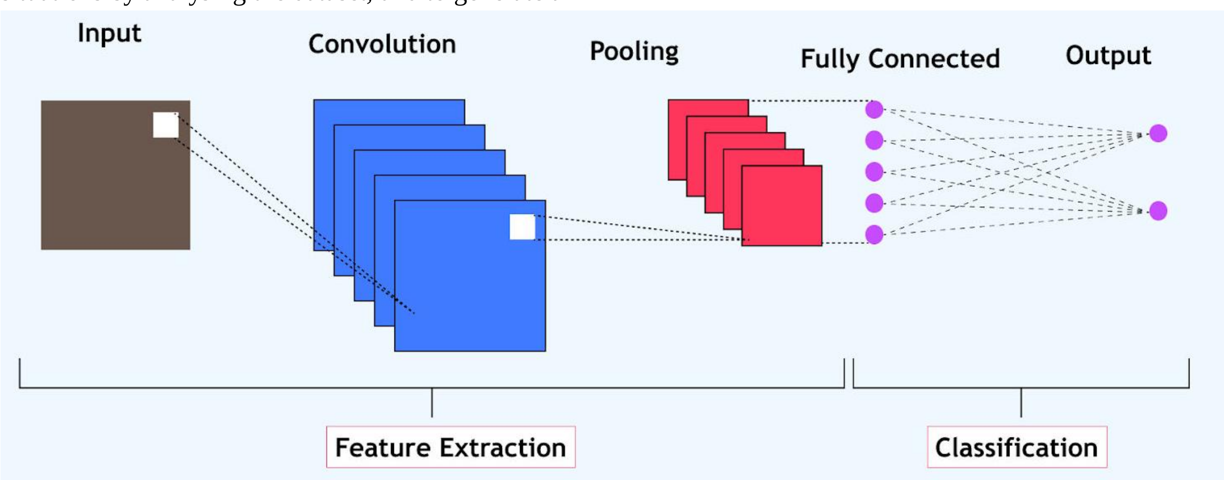

Figure 2: CNN Architecture (Gurucharan, 2024)

The 20images, each measuring 300 × 300 pixels, are used as input by the model. Multiple convolutional layers with different filter sizes and strides are used to extract the pertinent spatial characteristics from these images. Rectified Linear Unit (ReLU) (Agarap, 2018) activation functions are used to provide non-linearity and improve the model's ability to recognise and extract complex patterns. Additionally, a dropout layer is included to alleviate overfitting issues. Following flattening, the output of these layers is fed into several completely connected layers. The anticipated ego-vehicle distance value is produced by a linear activation function represented by the last layer. This model can be easily trained and deployed on several hardware platforms, including mobile devices, with over 5.9 million trainable parameters.

The suggested Convolutional Neural Network (CNN) model is made with a sequential architecture that is optimised for making predictions and extracting spatial data. The model starts with an input layer set up to receive 20 300×300 pixel image frames, which correspond to a structured input shape of (300, 300, 20). Five convolutional layers with different numbers of filters and kernel sizes make up the architecture's core; the first layer has 24 filters, while the subsequent layers have 64 filters. A rectified linear unit (ReLU) activation function is combined with each convolutional layer to add non-linearity and improve the model's capacity to recognise intricate patterns. Effective extraction of both fine-grained and large-scale spatial characteristics from the input images is ensured by the employment of different filter sizes and strides.

The model includes a dropout layer with a rate of 0.5 to combat overfitting. This layer randomly disconnects connections during training in order to encourage generalisation. Three fully connected layers with decreasing numbers of units (100, 50, and

10) receive the flattened output from the convolutional layers. ReLU activation is used in each layer to further improve the learnt features. The model can provide a continuous value that represents vehicle speed thanks to the last layer, which is a dense layer with a single unit and a linear activation function. With over 5.9 million trainable parameters, the CNN balances efficiency and complexity, making it suitable for hardware platforms or mobile devices and strong for predictive jobs. For real-world applications, our sturdy architecture maintains computing efficiency while guaranteeing high forecast accuracy.

1.4 Multiple Output CNN Model for Accident Prevention

The 1D signal for the ego-vehicle distance, which serves as the measurement input, is added to the network as an extra input. As a result, 20 road scene images and 20 distance data will be accepted by the updated network. The distance-measurement data that has been reshaped to preserve the same dimensionality and order is then concatenated with the flattened output following the convolutional layers from the first 20 image input. The same four dense layers from the original model come next. The last dense layer of the enlarged model will have three units (instead of one) since the model's output must take into consideration the additional projected data, the brake-pedal pressure, and the emergency-brake signal. In Figure 3, the network's structure is displayed.

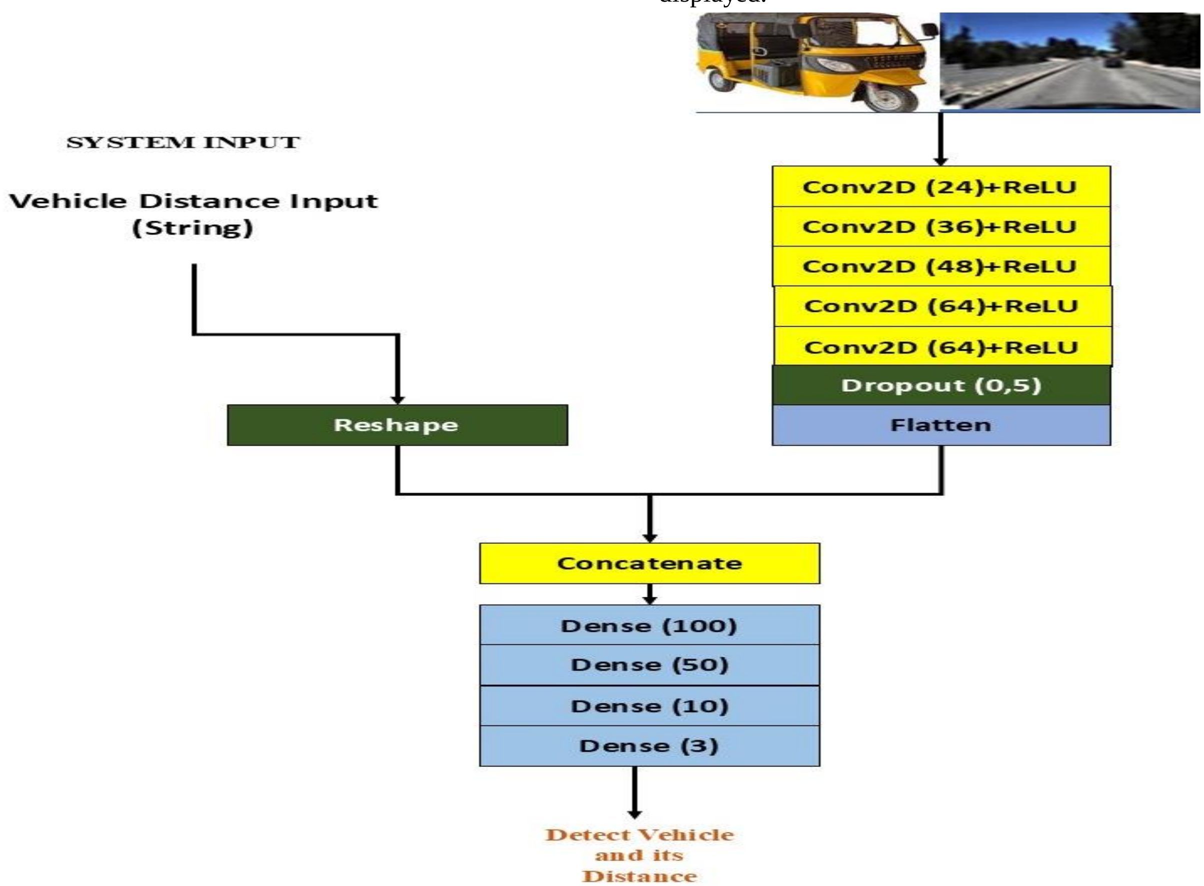

Figure 3: The Modified CNN Model for Vehicle Collision Prevention

A deep learning-based car safety system that combines numerical vehicle distance data and image-based inputs for braking signal prediction is depicted in Figure 3. Five convolutional layers with increasing filter sizes (24, 36, 48, 64, 64), ReLU activation, and a dropout layer (0.5) to prevent overfitting are used by the CNN branch to process visual inputs from road images. A flattening operation is then performed. At the same time, the vehicle distance input, which is given as a string, is reshaped to match the proportions of the CNN-extracted features. Following concatenation, the two feature sets are sent through four fully connected layers (100, 50, 10, and 3 neurones). The last layer then predicts three outputs: the distance to the vehicle, the brake signal, and the emergency braking signal. Real-time braking decisions for improved vehicle safety are made possible by this model's efficient integration of numerical and spatial data.

1.5 Training of the CNN Model

The Mean Squared Error (MSE) loss function, which quantifies the difference between the actual ground truth values and the predicted values, was used to train the CNN model that we demonstrated. By calculating the square of the differences between these values, the MSE loss function penalises bigger deviations more severely. The CNN optimises its settings to reduce the differences between its predictions and the real targets thanks to its robustness to errors. It offers an easily comprehensible indicator of the network's accuracy. We started the Adam optimiser (Kingma et al., 2014) with a learning rate of 0.001 and a decay of 0.0001 in order to use it during training. This helps with fine-tuning in later epochs and quick convergence in the early training phases. By preventing overshooting throughout the optimisation process, the learning rate decay, also known as scheduling, makes sure that the model keeps learning efficiently while training goes on.

2. SYSTEM IMPLEMENTATION

The models were implemented using image acquisition toolbox, data acquisition toolbox, machine learning and statistics toolbox, neural network toolbox and Simulink. The data acquisition and image acquisition toolbox were used to drive the data capturing process. The statistics toolbox was configured with the feature extraction techniques adopted. The neural network toolbox was configured with the accident prevention and control system developed. These toolboxes were used to implement the new system in Simulink environment and the performance evaluated.

3. RESULTS OF THE ACCIDENT PREVENTION AND CONTROL SYSTEM

Having tested, analysed and validated the result of the models developed for accident prevention and control model developed, it was used to develop an accident control system and evaluated as shown in the Figure 4;

Figure 4: Results of system integration with two vehicles

Figure 5: Result of the system testing with one threat vehicle

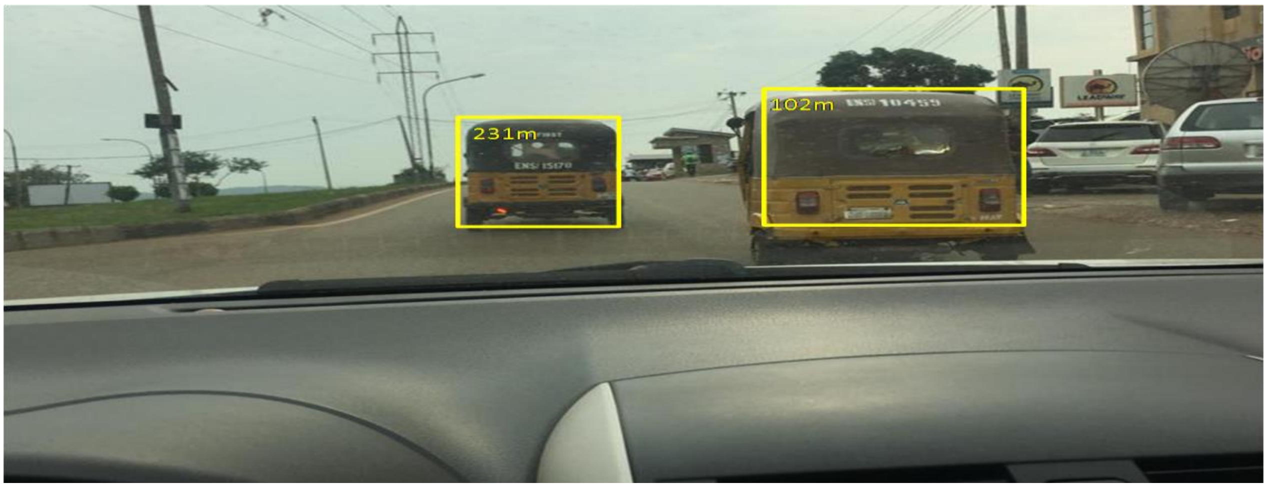

Figure 6: The result of the system integration

The Figure 4 presented the system integration result of the developed model for accident prevention and control. The result showed how the image sensor was able to capture the vehicle and classify as tricycle using the classification model, the proximity sensor was used to detect the distance of the vehicle from the main vehicle and then when it is less then 152m from the main vehicle, the sleep was controlled.

From the result it was observed that the accident detection model was able to classify one of the two tricycle distance as less than 152m in the front of the main vehicle, during translation and the speed was controlled. The next result presented anther system evaluation as in Figure 5;

The Figure 5 presented where the accident detection and control system were able to detect a tricycle which was 92meters in front of the main vehicle and then control to prevent accident. The next result presented the performance of the in another driver scenario and it was observed that potential accidents were detected and then controlled as shown in the Figure 6;

3.1 Comparative Analysis

In the comparative analysis, some of the sophisticated accident detection and control models developed over the years were analysed as shown in the Table 1;

Table 1 presented the comparative accuracy performance of the various accident detection and control model developed over the years. The percentage improvement from the best existing model is 5.8%. The results were analyzed using the graph in Figure 7;

Table 1: Comparative analysis

Authors	**Technique**	**Accuracy (%)**
Mehboobetal et al. (2016)	Fuzzy logic	87.0
Nancy et al. (2020)	Clustering technique	93.0
New system	FFNN	98.1
Nejdet and Abdulhamit (2014)	Agglomerative Hierachical clustering	90.2
Nejdet and Abdulhamit (2014)	K-Mean	79.0
Sreyan et al. (2019)	Convolutional neural network	95.0

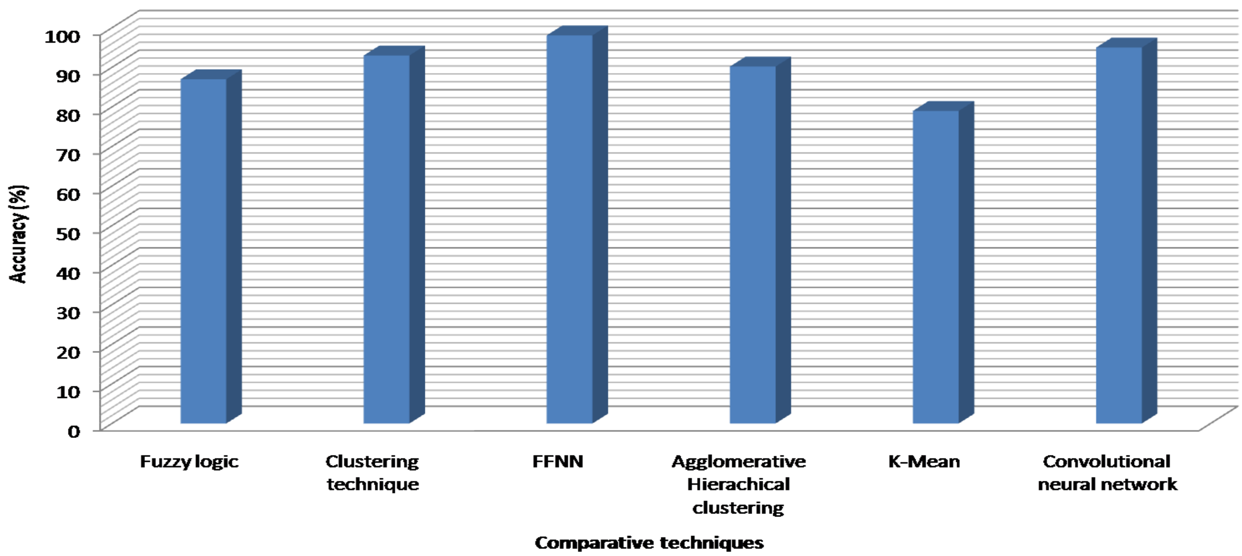

Figure 7: Comparative accuracy performance

The Figure 7 presented a comparative analysis of selected accident detection and control models developed over the years and the new system developed with the FFNN. The result showed that the FNN achieved better accuracy when compared with the others. The reason was due to the Hog feature extraction approach which was used to extract the required data or training with the FNN. Because enough and quality data was trained, the performance of the classification model was better. The next result presented the ROC performance in a comparative form as in the Table 2 while the graphical analysis was presented in the Figure 8;

Table 2: Comparative ROC performance

Author	**Techniques**	**ROC**
Bokaba et al. (2022)	Naïve Bayes	0.82
Bokaba et al. (2022)	Linear regression	0.41
Bokaba et al. (2022)	Support vector machine	0.16
Bokaba et al. (2022)	Adaboost	0.97
New system	FFNN	0.98

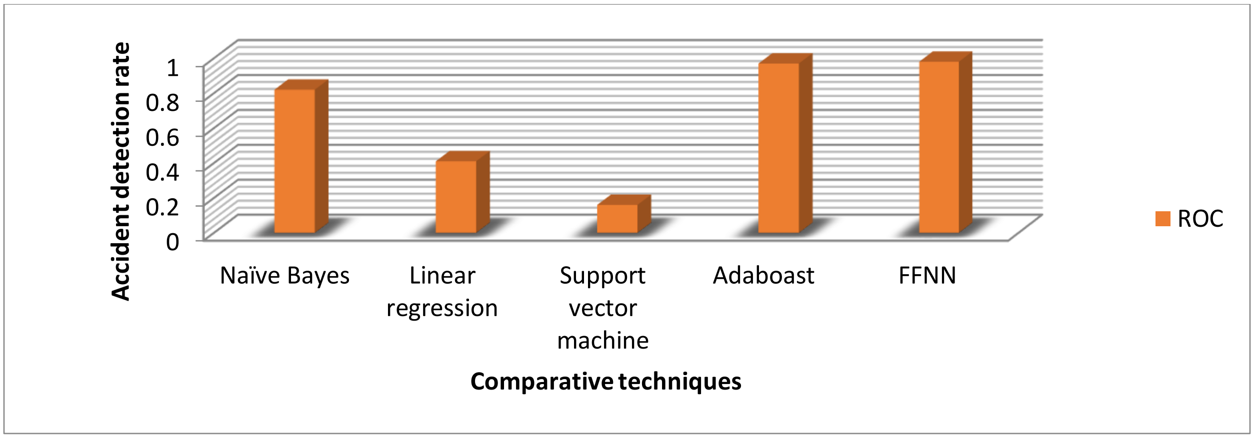

Figure 8: Comparative analysis of accident detection techniques

The Figure 8 presented a comparative performance of accident detection models. The result showed that the FNN achieved better detection performance when compared with existing system.

4. CONCLUSION

A Convolutional Neural Network (CNN) based system is developed for accident prevention and control to predict critical vehicle parameters like speed, brake pedal pressure and emergency brake signals. The initial model used a sequence of 20 images paired with the corresponding distance values to estimate the vehicle speed and the extended model predicted multiple outputs simultaneously by incorporating numerical distance data. With over 5.9 million trainable parameters, the CNN architecture achieved remarkable computing efficiency and accuracy by utilising sophisticated feature extraction techniques such as the Histogram of Gradient Features (HOG) and incorporating dropout layers to reduce overfitting.

The Honda Deep Drive (HDD) dataset and extra tricycle data that was gathered locally were used to test the system, which was put into practice in MATLAB's Simulink environment. This produced 21,300 samples. The system successfully detected dangerous conditions, such hard braking, and managed vehicle behaviour to avoid collisions, according to performance studies. The suggested model outperformed conventional CNNs and clustering strategies, with a detection accuracy of 98.1%. Its sturdy design, which integrates numerical and visual data, allows for real-time decision-making and guarantees increased safety in traffic situations. This study concludes by demonstrating the effectiveness of a CNN that has been optimised for real-time accident prevention and control. The model is a useful addition to intelligent car safety systems because of its exceptional accuracy and dependability as well as its capacity to process a variety of inputs. The study illustrates the significance of combining sophisticated deep learning architectures with high-quality data by outperforming state-of-the-art methods in terms of accuracy and ROC performance. The system's promise in autonomous and semi-autonomous driving technologies can be further advanced by future work that broadens its applicability to a wider range of situations and vehicle types.

REFERENCES

Aamir, M., Rahman, Z., Abro, W. A., Tahir, M., & Ahmed, S. M. (2019). An optimized architecture of image classification using convolutional neural network. International Journal of Image, Graphics & Signal Processing, 11(1), 30–39.

Agarap, A. F. (2018). Deep learning using rectified linear units (ReLU). arXiv preprint, arXiv:1803.08375.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded up robust features. Computer Vision – ECCV, 404–417. https://doi.org/10.1007/11744023_32

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., &Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. Sensors, 21(11), 3758. https://doi.org/10.3390/s21113758

Bokaba, T.,Doorsamy, W., & Paul, B.S. (2022) Comparative Study of Machine Learning Classifiers for Modelling Road Traffic

Accidents. Appl. Sci. 12, 828. https://doi.org/10.3390/app12020828

Chen, X., Gong, R. X., Xie, L. L., Xiang, S., Liu, C. L., & Pan, C. H. (2017). Building regional covariance descriptors for vehicle detection. IEEE Geoscience and Remote Sensing Letters, 14(4), 524–528. https://doi.org/10.1109/LGRS.2017.2658698

Gurucharan M.K., (2024) Basic CNN Architecture: A detailed explanation of the 5 layers in convolutional neural networks. https://www.upgrad.com/blog/basic-cnn-architecture/ Accessed February 2025

He, K., Gkioxari, G., Dollar, P., &Girshick, R. (2017). Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017, 2980–2988. https://doi.org/10.1109/ICCV.2017.322

Itu, R., &Danescu, R. (2024). Fully convolutional neural network for vehicle speed and emergency-brake prediction. Sensors, 24(1), 212. https://doi.org/10.3390/s24010212

Kim, J., Baek, J., & Kim, E. (2015). A novel on-road vehicle detection method using πHOG. IEEE Transactions on Intelligent Transportation Systems, 16(6), 3414–3429. https://doi.org/10.1109/TITS.2015.2443985

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint, arXiv:1412.6980.

Mantripragada, K., Trigo, F. C., & Martins, F. P. R. (2014). Vehicle tracking using features matching and Kalman filtering. ABCM Symposium Series in Mechatronics, 6, 1–8.

National Highway Traffic Safety Administration, United States Department of Transportation. (2023). Speeding: Risky driving. Retrieved from https://www.nhtsa.gov/risky-driving/speeding

Sadek, S., Al-Hamadi, A., Michaelis, B., & Sayed, U. (2010). Real-time automatic traffic accident recognition using HFG. 20th International Conference on Pattern Recognition (ICPR), 3348–3351. https://doi.org/10.1109/ICPR.2010.817

Saxe, J., & Berlin, K. (2015). Deep neural network-based malware detection using two-dimensional binary program features. Proceedings of the International Conference on Malicious and Unwanted Software, Fajardo, Puerto Rico, 20–22. https://doi.org/10.1109/MALWARE.2015.7413680

WHO. (2023). World Health Organization—Road traffic injuries. Retrieved from https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

Wu, C. C., & Weng, K. W. (2017). The detecting and tracking system for vehicles. 10th International Conference on Ubi-media Computing and Workshops (Ubi-Media), Pattaya, 1–5. https://doi.org/10.1109/Ubi-Media.2017.8058026