Original Research Article

Development of Deep Neural Network Model for the Prediction of Road Crashes in Real Time

Abstract

Road safety remains a global concern with the number of deaths and injury recorded from road traffic accidents estimated to be 1.5 million and 50 million respectively by 2025. Despite being predictable and largely preventable, the trend of road traffic crash is on the rise in Nigeria with an annual average of 33.7 deaths per 100,000 people. Proactive technique such as real time traffic and crash prediction has the potential to reduce the likelihood of crashes and to improve post-crash response. GoogleNet Convolutional Neural Network was developed in this study to classify road conditions and predict crashes along Ondo – Akure single carriage highway in Nigeria. Traffic flow relationships were established for the empirical data collected through video technique and compared to Green shields, Greenberg and Underwood models. The results were found generally satisfactory at an average coefficient of correlation of 0.96. The developed GoogleNet Convolutional network performed quite satisfactorily at predicting the probability of different traffic conditions – congested traffic (0.98), free-flowing traffic (0.64) and traffic crash (0.94). The developed algorithm can be integrated with traffic cameras and crowd-sourced images in areas that are not within the reach of surveillance cameras and sensors to report traffic condition in real time.
Keywords: Road Safety, Convolutional Neural Network, Traffic Flow, Deep Learning

1.0 Introduction

Road traffic crashes remains a major problem confronting transportation planners, policy makers and vehicle manufacturers all over the world (Ipindola, 2019). The statistics keeps growing at an alarming rate in spite of innovations and technological advancements in the transportation niche (NHTSA, 2020). WHO (2018) claims that approximately 1.3 million people die each year as a result of road traffic crashes. They went further to report that 93% of the world's fatalities on the roads occur in low- and middle-income countries, even though these countries have approximately 60% of the world's vehicles. Nigeria in context records an average of 230 crashes per 10,000 cars According to a yearly report published by Federal Road Safety Corps (FRSC, 2017). A large percentage of these crash cases are recorded on single carriage highways characterized by mixed traffic, speed violation, poor surface condition and weak lane discipline (Ipindola and Falana, 2019). Post-crash events are as important as pre-crash events in managing road safety (Zheng et al. 2014; Krivda et al., 2020). For instance, a significant percentage of fatality is recorded post-crash in Nigeria (Ipindola, 2019). Intelligent response has the potential to reduce post-crash effects such as shock waves, congestion and secondary crashes (Sayarshad, 2022). In a bid to solve the problem of delayed response, governments try to establish more response commands and units in strategic road traffic crash hotspots. This approach is obviously capital and human resource intensive and may not be able to compete favourably within government’s budget priorities. However, the unprecedented advancements in information technology (IT) coupled with improved computing power and internet of things (IoT) avails essential tools and techniques that can be leveraged to improve response through smart evacuation and avoidance of shockwave and other ripple effects. Moreso, first-hand crash features and characteristics can be useful for informed insurance claims (Baker, 2022).

To deal with the limitations of statistical methodologies, popular convolutional neural networks for object detection and classification such as AlexNets, GoogleNet and ResNet50 have been found effective in road safety applications (Sharma et al., 2018). These techniques have demonstrated high capacity to perform several image processing and computer vision applications in several industrial and scholastic studies in recent years (Zhou et al., 2014). The entire process first involves feature detection of images by selecting key points or forming a Grid over images, the choice made in order to speed up the process of detection. Then comes the stage of feature extraction for which SURF, a binary feature descriptor is employed. K-means clustering is then applied in order to quantize and make the bag of visual words. Every image, expressed as a histogram of visual words is fed to a supervised learning model, SVM for training. SVM is then tested for classification of images into respective classes. This technique is becoming more popular among scholars as evident in the ton of articles being published consistently (Gang and Xiaochi, 2012; Zhou et al., 2013). Prantly et al. (2017) developed a novel approach to detect features for 3D head modelling and animation based on curvature and its derived descriptors such as shape index, curvedness and Willmore energy. Zhou et al. (2012) presented a simple and effective scene classification approach based on the incorporation of a multi-resolution representation into a bag of feature model. They claimed that this proposed approach performed competitively with previous methods across all dataset.

Image processing is not an entirely new technique in the field of transportation. Several studies have employed image processing and deep learning algorithms to detect traffic signs (Yu et al., 2016; Rodriguez et al. 2022), pedestrians, road lanes, curves and other geometric features (Yadav, et al. 2022), and vehicular traffic conditions (Yuan et al., 2016) to name but a few. Yuan et al. (2016) proposed an unsupervised feature learning algorithm with encoded density information to classify congested scene. The study of (Zhang et al., 2018) is closely related to the aforementioned. They employed deep learning approach in detecting traffic accident from social media data. Vij and Aggarwal (2018), proposed a cost-effective approach to infer the traffic state of the road by analyzing the cumulative acoustic signal collected from the microphone sensor of user's smart phone to capture the distinctive characteristics of various traffic scenes, they explored two different types of features: Mel Frequency Cepstral Coefficients (MFCCs) and Wavelet Packet Transform (WPT). This study proposes a deep learning approach to predicting traffic incident images extracted from real time traffic flow by developing GoogleNet CNN algorithm in MATLAB programming environment.

Video survey was conducted for a period of three months on Ondo – Akure single carriage highway, the traffic features were extracted using image processing technique and afterwards classified in conjunction with crash data retrieved from FRSC Ondo command using GoogleNet Convolutional Neural Network technique in MATLAB programming environment.

2.0 Methodology

2.1 Data Collection

In a bid to achieve the aim of this research, video survey of traffic at the selected case study was conducted from September to December, 2021 as shown in Figure 1. The recorded video was imported into MATLAB programming environment and the features of interest extracted using suitable image processing algorithm.

Figure 1: Traffic Survey being conducted at Case Study

2.2 Traffic Flow Relationship

The traffic features of interest extracted from the video survey include: speed, volume, density and percentage of heavy vehicles. Regression curves were fitted to the empirical data, the fundamental speed-flow-density relationship was established and the result compared with conventional Greenshields, Greenberg and Underwood models.

2.3 Data Reprocessing

In order for the GoogleNet CNN technique to be effective, the extracted images were cropped to ensure that the majority of their areas were occupied by the subject of category. Next, an array of image sets based on accident, congested traffic and free-flowing traffic were constructed. To manage the data image set, class was employed to operate on image file locations. Each element of the image set variable now contains images associated with the particular category. Two thousand images were included in each category and were divided randomly in proportion of 70% for training and 30% for validation as presented in Table 1 to avoid biased results.

Table 1: Training and Validation Dataset for each Image Category

Image Category	Training dataset	Validation dataset
Accident	1400	600
Congested traffic	1400	600
Free-flowing traffic	1400	600

2.4 Development of GoogleNet CNN Algorithm

An algorithm was developed to train and validate the traffic incident and condition images extracted from video survey as shown in Figure 2 below.

```
Dataset = imageDatastore('dataset', 'IncludeSubfolders', true, 'LabelSource', 'foldernames');
[Training_Dataset, Validation_Dataset] = splitEachLabel(Dataset, 7.0);
net = googlenet;
analyzeNetwork(net)

Input_Layer_Size = net.Layers(1).InputSize(1:2);
Resized_Training_Image = augmentedImageDatastore(Input_Layer_Size, Training_Dataset);
Resized_Validation_Image = augmentedImageDatastore(Input_Layer_Size, Validation_Dataset);

Feature_Learner = net.Layers(142);
Output_Classifier = net.Layers(144);

Number_of_Classes = numel(categories(Training_Dataset.Labels));

New_Feature_Learner = fullyConnectedLayer(Number_of_Classes,...
    'Name', 'Traffic Feature Learner',...
    'WeightLearnRateFactor', 10, ...
    'BiasLearnRateFactor', 10);

New_Classifier_Layer = classificationLayer('Name', 'Traffic Classifier');

Layer_Graph = layerGraph(net);

New_Layer_Graph = replaceLayer(Layer_Graph, Feature_Learner.Name, New_Feature_Learner);
New_Layer_Graph = replaceLayer(New_Layer_Graph, Output_Classifier.Name, New_Classifier_Layer);

analyzeNetwork(New_Layer_Graph)

size_of_Minibatch = 5;
Validation_Frequency = floor(numel(Resized_Training_Image.Files)/size_of_Minibatch);
Training_Options = trainingOptions('sgdm',...
    'MiniBatchSize', size_of_Minibatch,...
    'MaxEpochs', 6,...
    'InitialLearnRate', 3e-4,...
    'Shuffle', 'every-epoch',...
    'ValidationData', Resized_Validation_Image,...
    'ValidationFrequency', Validation_Frequency,...
    'Verbose', false,...
    'Plots', 'training-progress');

net = trainNetwork(Resized_Training_Image, New_Layer_Graph, Training_Options);
```

Figure 2: Algorithm for training GoogleNet CNN for Traffic Incident Classification

2.5 GoogleNet CNN Learning Architecture

The GoogleNet CNN developed for this study comprise 144 layers –1 input, 142 hidden and 1 output layer as presented in Figure 3. The input size comprises 224 by 224 by 3 i.e., 224 by 224 pixel image having 3 channels. To train the network, a minibatch of size 5 was selected.

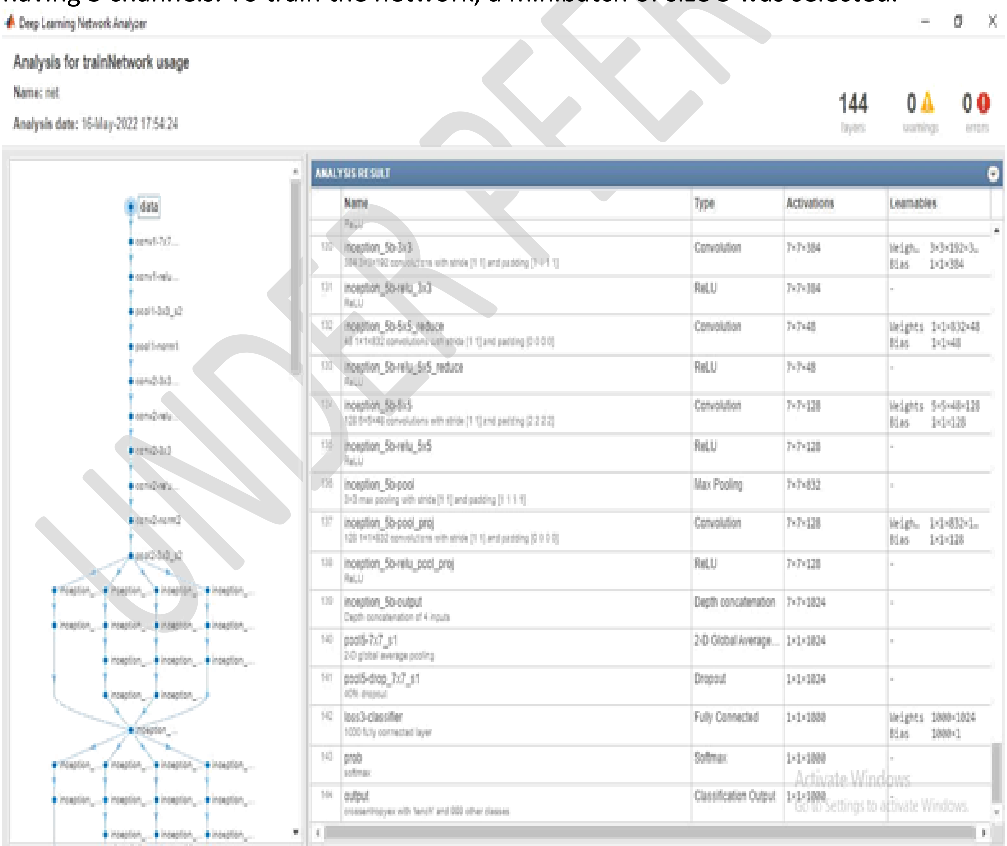

Figure 3: GoogleNet CNN Architecture in MATLAB Deep Learning Analyzer

2.6 Testing the GoogleNet Convolutional Neural Network

A function was developed to test the accuracy of the developed GoogleNet CNN at predicting traffic incidents and conditions as shown in Figure 4.

```
function test_network(net, image)
I = imread(image);

R = imresize(I, [224, 224]);
[Label, Probability] = classify(net, R);
figure;
imshow(R);
title({char(Label), num2str(max(Probability), 6)})
end
```

Figure 4: Algorithm for Testing GoogleNet CNN

3.0 Results and Discussion

3.1 Traffic Flow Relationship

The fundamental relationship between speed-flow-density was explored for the empirical data extracted from traffic video using image processing algorithm and an empirical regression curve was fitted to the resulting features of speed, density and flow as presented in Figures 5. The empirical data was also fitted to Greenshields, Greenberg and Underwood models and the results compared in Table 2. Greenshields model was found to fit the empirical data the most while sharing the same optimum speed, jam density and free-flow speed values with the empirical model. The maximum flow value of 3758km/h was recorded by the empirical model followed closely by Greenshields with 3712km/h. The maximum flow values of Greenberg and Underwood models were far lower than the values obtained by the empirical and Greenshields model. This may be due in part to the logarithmic and exponential relationships between the speed and the density of Greenberg and Underwood models respectively. The free flow speed of Greenberg model (112 km/h) was found to be way higher than the 91km/h obtained from the other three models. Caution must be exercised in applying these conventional models to solving traffic problems. The jam densities of the empirical, Greenshields and Greenberg models were in close ranges while that of Underwood tended to infinity. Considering the coefficient of correlation of the four models under comparison, it can be generally concluded that the models performed satisfactorily at fitting the empirical data.

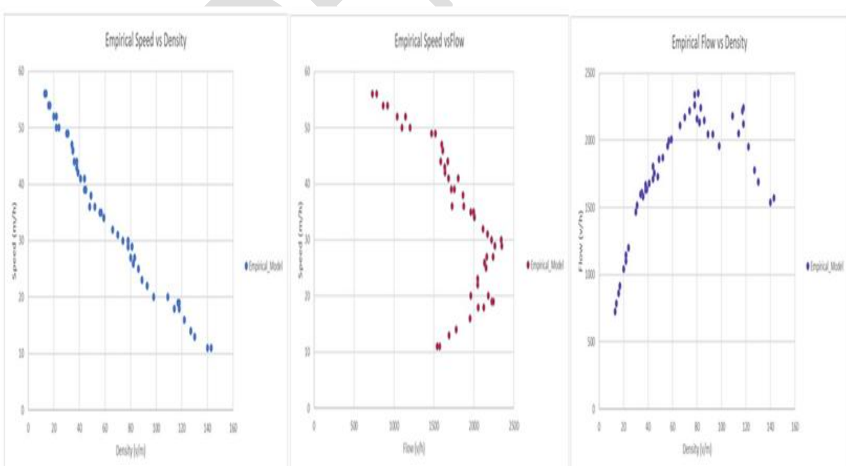

Figure 5: Speed-Flow-Density Curves for the Empirical Data

Table 2: Comparison of Empirical and other Conventional Models

Model	Optimum Speed	Optimum density	Maximum flow	Jam density	Free-flow speed	Coef. Of Cor. (R^2)
Empirical Model **($y = -0.3484x + 56.948$)**	46	131	3758	360	91	0.97
Greenshields Model	46	128	3712	312	91	1.00
Greenberg Model **($y = -29\ln(x) + 147.72$)**	46	72	2782	342	112	0.95
Underwood Model **($y = 57e^{-0.012x}$)**	32	138	2747	Infinity	91	0.95

3.2 GoogleNet CNN Model Summary

The accuracy of the trained image category classifier based on the training dataset was found to be 0.81 while its accuracy based on validation dataset was 0.66 as showed in Figure 6. It can be observed from the training progress graph that the developed architecture improved in accuracy as the iteration progressed. The learning rate was set at 3e-4 and 6 Epoch were achieved with the frequency of iteration being 4. The training cycle comprise; 6 Epoch, 24 iterations at 4 iterations per Epoch. The evaluation produced a satisfactory confusion matrix as shown in Tables 3 and 4.

Figure 6: Accuracy vs Iteration Result of GoogleNet CNN Classifier

Table 3: Confusion Matrix for the Training Sets

	Predicted		
Known	Accident	Congested traffic	Free-flow traffic
Accident	0.82	0.10	0.0
Congested traffic	0.0	1.0	0.0
Free-flow traffic	0.1	0.0	1.0

Table 4: Confusion Matrix for the Validation Sets

Known	Predicted		
	Accident	Congested traffic	Free-flow traffic
Accident	0.66	0.14	0.11
Congested traffic	0.06	0.79	0.08
Free-flow traffic	0.1	0.0	0.70

3.3 Validation of Developed CNN Model

The developed GoogleNet CNN model performed quite satisfactorily at classifying different traffic condition and incident images. Figure 7 below shows the deep learning classifier predicting the probability of different traffic conditions –congested traffic (0.98), free-flowing traffic (0.64) and traffic crash (0.94). The algorithm can be integrated with traffic cameras and crowd-sourced images in areas that are not within the reach of surveillance cameras and sensors to report traffic condition in real time.

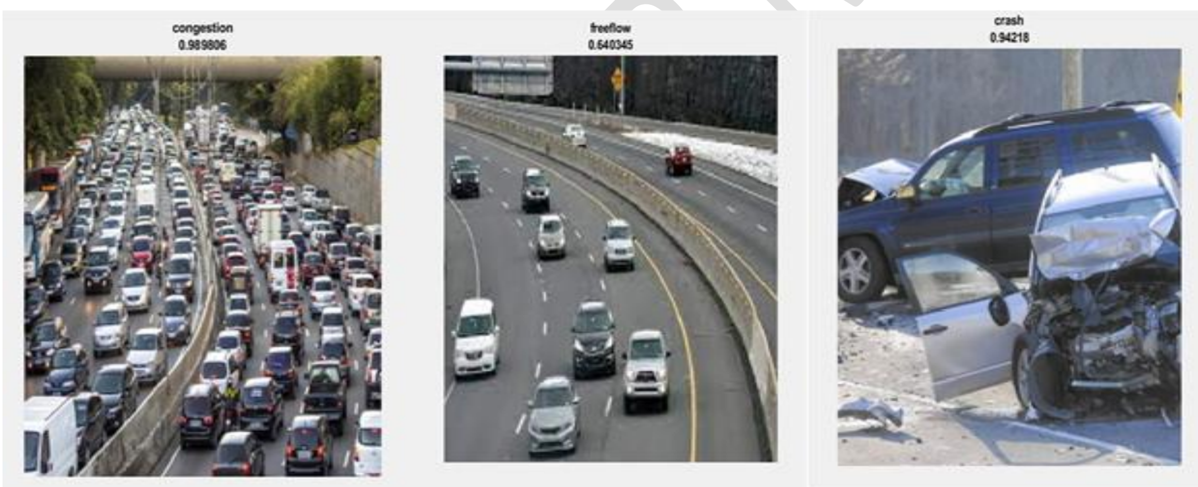

Figure 7: Result of GoogleNet CNN model at predicting Traffic Condition using Test Dataset

4.0 Conclusion

GoogleNet Convolutional Neural Network Model have been employed in this study to classify and predict road traffic condition based on image classification. The volume, variety and veracity of crowd-sourced dataset coupled with improved computing power and recent developments in machine learning techniques have proven to be the game changer in the field of computer vision applications in road traffic safety and emergency management. The performance of the classifier developed in this study was satisfactory, though there is room for improvement. Future research will focus on merging mined texts, location information and images in order to improve on the classification of road traffic conditions and accidents towards effective and intelligent emergency response.

References

1. Baker. C. (2022). Car Insurance Claims. wallet Hub, retrieved from https://wallethub.com/edu/ci/car- insurance-claims/12882 on the 10th of May, 2022.
2. Deng, L., Platt, J. (2014). Ensemble Deep Learning for Speech Recognition. *In Proc. Interspeech,* 2014.
3. Gang, L. and Xiaochi, w. (2012). Improved Bags-of-Words Algorithm for Scene Recognition. Physics Procedia 24 (2012) 1255 –1261.
4. Ipindola O. O. and Falana J. N. (2019). Traffic Stream Relationships of Two-Lane Highways: A Case of Akure-Ondo Road in Southwest Nigeria, *International Journal of Advanced Engineering, Management and Science,* 5(1), 87-93.
5. Ipindola, O. O. (2019). Multinomial Logit Approach to Modelling Crash Severity on Two Lane Highways in Nigeria. *Proceedings of the 1st International Conference of Engineering and Environmental Sciences Osun State University*, pp. 244-256.
6. Krivda, V. Petru, J., Macha, D. Plocova, K. and Fibich, D. (2020). An Analysis of Traffic Conflicts as a Tool for Sustainable Road Transport. Sustainability 2020, 12, 7198.
7. NHTSA (2020). NHTSA Releases 2020 Traffic Crash Data, Showing highest number of fatalities and highest fatality rate since 2007.
8. Rodríguez, R. C., Carlos, C. M., Villegas, O. V., Sánchez, V. G. and Domínguez, H. O., (2018). Mexican traffic sign detection and classification using deep learning. Expert Systems with Applications, Volume 202, 15 September 2022, 117247.
9. Sayarshad, H. R. (2022). Designing an intelligent emergency response system to minimize the impacts of traffic incidents: a new approximation queuing model. International Journal of Urban Sciences, DOI: 10.1080/12265934.2022.2044890.
10. Vij, D. and Aggarwal, N. 2018. Smartphone based traffic state detection using acoustic analysis and crowdsourcing. Applied Acoustics, 138, 80-91.
11. WHO (2018). Road traffic injuries. Retrieved from https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries on the 7th of May, 2022.
12. Yadav, M. Khan, P. and Singh, A. K. (2022). Identification of Pole-like Objects from Mobile Laser Scanning Data of Urban Roadway Scene. Remote Sensing Applications: Society and Environment, Volume 26, April 2022, 100765.
13. Yuan, Y. Wan, J. and Wang, Q. (2016). Congested scene classification via efficient unsupervised feature learning and density estimation. Pattern Recognition, 56 (2016), 159-169.
14. Yu, R., Abdel-Aty, M. A., Ahmed, M. M., Wang, X. (2014). Utilizing Microscopic Traffic and Weather Data to Analyze Real-Time Crash Patterns in the Context of Active Traffic Management. *IEEE Transactions on Intelligent Transportation Systems*, 15(1), 205-213.
15. Zhang, Z., He, Q. Gao, J. Ni, M. (2018). A deep learning approach for detecting traffic accidents from social media data. Transportation Research Part C, 86(2018), 580 –596.
16. Zheng, L., Ismail, K. and Meng, X. (2014). Traffic conflict techniques for road safety analysis: open questions and some insights. Canada Journal of Civil Engineering, 41(7), 213 –220.
17. Zhou, L., Zhou, Z. and Hu, D. (2013). Scene classification using a Multi-Resolution Bag-of-Features Model. Pattern Recognition 46 (2013) 424–431.