Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

Indian Road Conditions and Accident Risk Predictions using Deep Learning Approach – A Review

Partheeban Pachaivannan
Civil Engineering
Chennai Institute of Technology
Chennai, India
principal@citchennai.net

Rani Hemamalini Ranganathan
Electrical and Electronics Engineering
St. Peter's Institute of Higher Education and Research
Chennai, India
ranihema@yahoo.com

P. Navin Elamparithi
Computer Science and Engineering
National Institute of Technology
Tiruchirapalli, India
parithi1999@gmail.com

R. Dhanagopal
Electronics and Communications Engineering
Chennai Institute of Technology
Chennai, India
dhanagopalr@citchennai.net

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9316128

Abstract— **Indian road conditions are very concerning as they are traffic prone in many locations. Further, the road surface conditions are not uniform in several places which may lead to a sudden impact on the vehicle dwellers on the highways. Furthermore, the driver may be the cause of accidents due to ignorance and improper focus on the road during driving. Statistical sources show that there are many reasons regarding the risk of road accidents to consider while calculating the reasons and the risk of road accidents. Various factors lead to road accidents and hence a strong methodology and the support of computer-aided support is required to solve such critical problems. The current work deals with the application of the deep learning approach to predict critical road accidents.**

Keywords—road accident, deep learning

I. INTRODUCTION

India is a developing with a burgeoning population of about 1.35 billion people. With the number of registered vehicles going from about 230 million in 2016 to 232 million in 2019 and the numbers steadily increasing. The percentage of households owning a car went from 5% in 2011 to 11% in 2016. There are more and more vehicles entering the roads of India and most of these roads are poorly laid and seldom maintained. These conditions are exacerbated by heavy vehicles and weather conditions that cause the roads to be filled with potholes and many irregular deformities. The state of Tamil Nadu has recorded the highest number of traffic collisions for more than a decade and the capital city of Chennai has recorded the most collisions of any city in India. Chennai has had an alarming reputation for having one of the highest mortality rates from traffic crashes worldwide since the 1960s, at a period when the number of cars in the region was a tiny proportion of those of major megacities around the world, such as New York and Tokyo. New Delhi, the capital of India, has a 40-fold difference in the number of road accidents compared to that of London. Data from 2015 indicated that one person dies from traffic collisions every four minutes according to an NGO, 'Indians for Road Safety'. The numbers reported might be lower than actual casualties as many accidents go unreported. The World Health Organization (WHO) released the 'GlobStatus Report on Road Safety' which described the main causes of traffic accidents as speeding, driving under the influence of alcohol, and not wearing seat belts and helmets. Refusal to yield to oncoming traffic or establish a lane when making a turn is the primary cause of road accidents on national highways. The users of bikes and three-wheelers were the second largest group of fatalities according by the report. All of these factors, along with a general disregard for traffic rules, lead to one of the highest traffic accident numbers and casualties in the world.

II. DEEP LEARNING & DEEP LEARNING ARCHITECTURES

Deep learning is a subset of machine learning methods based on representation learning and artificial neural networks. Learning is of 3 types, namely, unsupervised, supervised or semi-supervised. The inspiration for artificial neural networks (ANNs) is the propagation of coordination nodes and transmission of information in biological systems. Therefore, ANNs are an extension from brains or to be precise Biological Neural Networks (BNNs).

Most deep learning models stem from the machine learning algorithm known as Artificial Neural Networks. To be more specific, the subset of ANNs known as Convolutional Neural Networks (CNNs) form the basis of most deep learning models. Deep learning is hence also referred to as Deep Neural Networking. The universal approximation theorem or probabilistic inference is used to interpret Deep Neural Networks (DNNs). The universal approximation theorem involves the ability of a network with expandable depth but a bounded width. The probabilistic interpretation involves inference, as well as teaching and testing modeling principles related to fitting and generalization, respectively.

To understand deep learning, imagine a kid whose first term is 'dog'. The baby knows what a dog is, not by pointing things and saying the word dog. The parent says, "Yes, that is a dog," or, "No, that is not a dog." As the child begins to point at things, he is more conscious of the characteristics that all dogs provide. What the child does, without realizing it, is to simplify the complicated complexity of the idea of a dog by creating a hierarchy in which each level of complexity is generated with the information that has been obtained from the previous layer

of the hierarchy. Similarly, a deep learning model is trained to analyse and decide/predict any given situation.

A child might be able to point out or identify a 'dog' within a few weeks or months, to train a deep learning model to identify a 'dog' with acceptable levels of accuracy takes an incredible amount of resources. Resources that include millions of pictures of 'dogs' to function as training data and the processing power required to process these pictures in this situation. Until a few years ago where the era of big data and cloud computing made it possible for programmers and researchers to access such data, it was impossible to create such elaborate and complex statistical models based on deep learning.

III. Critical Review of Literature

The most basic algorithm applied to the Indian accident dataset was done by (Rana *et al.*,)[1] as they attempted to predict road accidents in the Indian state of Maharashtra using 15 different variables in a logistic regression algorithm that produced an 'yes' or 'no' answer to whether an accident could take place using the aforementioned variables. Lee *et al.*, [2] evaluated three models of machine learning to predict traffic accident severity in the rainy season. They applied random forest algorithm, ANN and decision tree algorithm to precipitation data, traffic accident data and road geometry data for the Naebu Expressway for a span of nine years. Roop Kumar and Ramamurthy [3] applied ANN and decision tree algorithm to identify the attributes that have a severe impact on the accident severity label in Indian roads and defined four major factors that lead to accidents namely – driver, vehicle, road and environment. Pradhan *et al.*, [4] investigated the applications of deep learning in prediction of severity of traffic accidents by applying three architectures. These were Convolutional Neural Networks, Recurrent Neural Networks and simple feedforward Neural Networks. The results explained the accuracy of the performance of the three architectures and the best batch size and dropout probabilities as well.

Zhao *et al.*, [5] attempted a novel way of detecting accidents by way of social media. The findings revealed that paired tokens collected the association rules attributable to crash-related tweets as well as further improved the accuracy of road accident identification. Long Short-Term Memory (LSTM) and Deep Belief Network (DBN) were investigated and implemented on the extracted token. The supervised Latent Dirichlet allocation (sLDA) and Support Vector Machines (SVMs) were outperformed by the classification results obtained from the DBN. Formosa, Quddus, Ison, Abdel-Aty, and Yuan [6] predicted real time traffic conflicts using deep learning. Conflict data was generated using a Regional – CNN and the data was integrated and developed into a series of DNN models that predicted traffic conflicts. The time to collision (TTC) was found to be dependent on speed, weather and traffic density and the best DNN model employed was found to have 94% accuracy.

Zhao *et al.*, [7] proposed a deep learning algorithm to run based on the data collected from the Vehicular Ad-hoc Network (VANET). The algorithm was based on Convolutional Neural Networks and the extraction of different features were done by different convolution kernels. This led to new variables which were input into the established training model for training and testing. The results showed that in CNNs have lower loss and higher prediction accuracy comparison to traditional Back Propagation (BP) Neural Networks. Mohamed [8] predicted the causes of road accidents by means of a multi – class Support Vector Machine (SVM) model. Data was collected and data mining methods were used to define the causes of road accidents.

Chong *et al.,* [9] applied four different machine learning paradigms to analyse traffic accidents. The models used are decision trees, support vector machines, hybrid learning approaches and a concurrent hybrid model made up of both neural networks and decision trees. Results showed that the individual approaches were outperformed by the hybrid decision tree-neural network. Dong *et al.,* [10] researched an improved deep learning model for prediction of traffic crashes. The model is made up of 2 modules, a supervised module that performed traffic crash prediction and an unsupervised module that identified the functional network between the variables and the features. The supervised module in the regression layer was the multivariate negative binomial module.

Manchanda *et al.,* [11] utilised a Hybrid DNN that forecast traffic conditions on roads using CNN and predicted road accident statistics for those particular roads at a particular time. Results showed better results for prediction of traffic conditions and analysis of road accidents in comparison to other standard methods. Singh *et al.,* [12] investigated the usage of M5 model tree and RENB model for prediction of accidents in non-urban sections of the highways of India. In accordance to the 15 accident parameters both models had similar results in terms of RMSE values and correlation coefficient. The M5 model tree offered simple linear equations that are simple to understand and offer greater analysis, suggesting that this model can be easily used as an alternative to the RENB method if the primary goal is to forecast traffic accidents. The M5 model tree also revealed that the sensitivity analysis findings reflected the physical conditions of the roads.

Suganya and Vijayarani [13] investigated the usage of DNN and Gene Expression Programming (GEP) models to predict traffic accident frequency. The results obtained were compared to the standard Random Effect Negative Binomial (RENB) model. The DNN model had outperformed the RENB and GEP models and the RENB model was also outperformed by the GEP model in regards with all performance indicators. Singh *et al.,* [14] analysed road accidents in India using various classification algorithms such as logistic regression, linear regression, SVM, decision tree, Random Forest, KNN, Naïve Bayes and gradient boosting algorithm. Accuracy, error rate and execution time were used to measure to performance and the results indicated that KNN had better performance in comparison to the algorithms.

Vasavi [15] attempted to extract hidden patterns from road accidents and hence, obtain common attributes that may lead to accidents. Data visualization was done with the use of density histograms and the results obtained show that the machine learning techniques utilised were able to extract hidden patterns from the data. Ramani and Selvaraj [16] identified the set of factors that led to fatal road accidents in the roads around Coimbatore, Tamilnadu, India and the ones in the USA and the UK. They have applied Voting Algorithm for Aggregated Feature Selection to understand the severity of accidents and hence identify the factors that lead to fatal road accidents. Kumar and Toshniwal [17] analysed the severity of powered two-wheeler (PTW) accidents in the state of Uttarakhand, India. Three classification algorithms namely, decision tree (CART), Naïve Bayes and Support vector machine were applied on the PTW accidents dataset. Results showed that among the 13 districts of Uttarakhand, different combinations of factors were associated to the severity of the PTW accidents in different districts. Sharma *et al.,* [18] created a traffic accident prediction model using Support Vector Machines with Gaussian Kernel. The results obtained showed that the model had higher accuracy in comparison to the traditional Multilayer Perceptron (MLP) model.

IV. Need for Application of Deep Learning in Chennai City

Statistics published on road traffic accidents in 2018 by the Indian Ministry of Road Transport and Highways indicated that Chennai had the highest number of road accidents with 7,850 which resulted in the deaths of at least 1,260 people and left 7,438 injured as well. This is a distinction that Chennai can do without. The capital city, Delhi was second with 6,515 accidents, however, it topped the list of deaths with 1,690 people dead because of road accidents. The number of injuries from road accidents rose from 7,257 in 2017 to 7,580 in 2018 in Chennai. Delhi, Bengaluru, Bhopal and Indore preceded the state capital between cities where the number of injuries increased. Across the world, a total of 151,417 deaths were attributed to road accidents in 2018. This is a 2.3 per cent increase from the numbers in 2017. Approximately 85 per cent of accident-related deaths are reported by people in the 18-60 age range. Road injuries not only cause significant damage to the victim's families, but also cause a big economic loss to the region.

From the statistics given in Fig 1, the number of people involved in the accident is almost higher than the number of injuries. In the event of fatal, grievous and major incidents, the number of victims injured is much higher than the number of accidents. This data is necessary to be processed and analysed so that it can help decision makers put in place and create tailored measures to tackle the challenge of increasing road injuries and related casualties. To analyse this data, better techniques are required to provide better and more accurate models to road safety. These are the techniques and ideas listed above in the literature that were reviewed above and the cutting edge deep learning paradigms will lead way into a brave new world, where, the life without fear and risk of the unpredictable events known as traffic accidents.

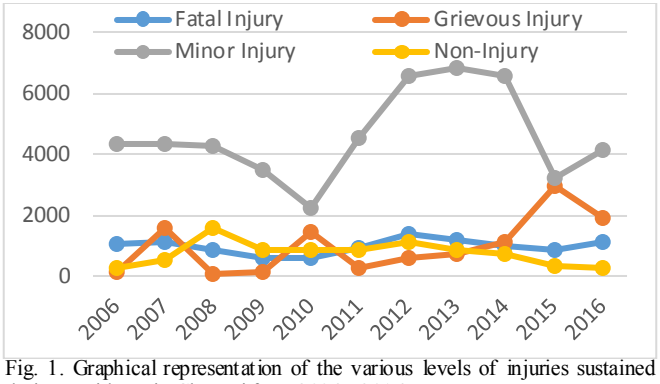

Fig. 1. Graphical representation of the various levels of injuries sustained during accidents in Chennai from 2006 – 2016

V. Conclusion

Indian roads have a reputation of being highly dangerous and everyday various accidents can be seen on the news. With over 5.9 million kms of surfaced roads, India has the second largest road network in the world behind the USA. With roads varying from huge 8 lane expressways to single lane rural roads, there are roads at some of the highest altitudes and densely populated cities in the world. All these various conditions mean that the factors influencing road accidents vary from location to location. For the state of Tamilnadu, especially the city of Chennai is notorious for the extremely high number of fatal road accidents. The data on these accidents especially the factors affecting the accidents and their severity is collected by the Chennai City Traffic Police Commissioner Office. As the country with the second highest population in the world, the need for the latest innovations in data analysis and deep learning models need to be applied in order to define and derive the factors and causes of accidents in India as well prevent such accidents from taking place. The dataset in hand needs to be analysed and modern methods need to be applied in order to implement a proactive system that strives to prevent accidents and reduce the risk of fatal accidents as well as the severity of accidents as well. The population and the vehicles on road need to be safeguarded as they are the lifeblood of the nation and the key to unlocking economic growth and maximising the potential of the country.

References

[1] V. Rana, H. Joshi, D. Parmar, P. Jadhav, and M. Kanojiya, "Road Accident Prediction using Machine Learning Algorithm," *Int. Res. J. Eng. Technol.*, vol. 3494, pp. 3494–3496, 2008, [Online]. Available: www.irjet.net.

[2] J. Lee, T. Yoon, S. Kwon, and J. Lee, "Model evaluation for forecasting traffic accident severity in rainy seasons using machine learning algorithms: Seoul city study," *Appl. Sci.*, vol. 10, no. 1, 2020, doi: 10.3390/app10010129.

[3] R. Roop Kumar and B. Ramamurthy, "Data Analysis in Road Accidents Using ANN and Decision Tree," *Int. J. Civ. Eng. Technol.*, vol. 9, no. 4, pp. 214–221, 2018, doi: IJCIET_09_04_023.

[4] B. Pradhan, M. I. Sameen, H. Z. M. Shafri, and H. Bin Hamid, "Applications of Deep Learning in Severity Prediction of Traffic Accidents," no. January, pp. 793–808, 2019, doi: 10.1007/978-981-10-8016-6.

[5] H. Zhao *et al.*, "A deep learning approach for detecting traffic accidents from social media data," *Transp. Res. Part C Emerg. Technol.*, vol. 86, no. 1, pp. 580–596, 2018, doi: 10.1016/j.trc.2017.11.027.

[6] N. Formosa, M. Quddus, S. Ison, M. Abdel-Aty, and J. Yuan, "Predicting real-time traffic conflicts using deep learning," *Accident Analysis and Prevention*, vol. 136. 2020, doi:

10.1016/j.aap.2019.105429.

[7] H. Zhao, H. Cheng, T. Mao, and C. He, "Research on Traffic Accident Prediction Model Based on Convolutional Neural Networks in VANET," *2019 2nd International Conference on Artificial Intelligence and Big Data, ICAIBD 2019*. pp. 79–84, 2019, doi: 10.1109/ICAIBD.2019.8837020.

[8] Chen, Joy Iong Zong, and S. Smys. "Social Multimedia Security and Suspicious Activity Detection in SDN using Hybrid Deep Learning Technique." *Journal of Information Technology* 2, no. 02 (2020): 108-115.

[9] M. Chong, A. Abraham, and M. Paprzycki, "Traffic accident analysis using machine learning paradigms," *Inform.*, vol. 29, no. 1, pp. 89–98, 2005, doi: 10.31449/inf.v29i1.21.

[10] C. Dong, C. Shao, J. Li, and Z. Xiong, "An Improved Deep Learning Model for Traffic Crash Prediction," *J. Adv. Transp.*, vol. 2018, 2018, doi: 10.1155/2018/3869106.

[11] C. Manchanda, R. Rathi, and N. Sharma, "Traffic Density Investigation & Road Accident Analysis in India using Deep Learning," in *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 2019, pp. 501–506, doi: 10.1109/ICCCIS48478.2019.8974528.

[12] G. Singh, S. N. Sachdeva, and M. Pal, "M5 model tree based predictive modeling of road accidents on non-urban sections of highways in India," *Accid. Anal. Prev.*, vol. 96, pp. 108–117, Nov. 2016, doi: 10.1016/j.aap.2016.08.004.

[13] G. Singh, M. Pal, Y. Yadav, and T. Singla, "Deep neural network-based predictive modeling of road accidents," *Neural Comput. Appl.*, vol. 8, 2020, doi: 10.1007/s00521-019-04695-8.

[14] E. Suganya and S. Vijayarani, "Analysis of road accidents in India using data mining classification algorithms," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, 2017, pp. 1122–1126, doi: 10.1109/ICICI.2017.8365315.

[15] S. Vasavi, "A Survey on Extracting Hidden Patterns within Raoad Accident Data using Machine Learning Techniques," *Commun. Appl. Electron.*, vol. 6, no. 4, pp. 1–6, 2016, doi: 10.5120/cae2016652455.

[16] R. G. Ramani and S. Selvaraj, "A Pragmatic Approach for Refined Feature Selection for the Prediction of Road Accident Severity," *Stud. Informatics Control*, vol. 23, no. 1, pp. 41–52, Mar. 2014, doi: 10.24846/v23i1y201405.

[17] S. Kumar and D. Toshniwal, "Severity analysis of powered two wheeler traffic accidents in Uttarakhand, India," *Eur. Transp. Res. Rev.*, vol. 9, no. 2, p. 24, Jun. 2017, doi: 10.1007/s12544-017-0242-z.

[18] B. Sharma, V. K. Katiyar, and K. Kumar, "Traffic Accident Prediction Model Using Support Vector Machines with Gaussian Kernel," 2016, pp. 1–10.