

# PREDICTING ROAD ACCIDENT RISK USING GOOGLE MAPS IMAGES AND A CONVOLUTIONAL NEURAL NETWORK

Aarya Agarwal

Westwood High School, Austin, USA

## **ABSTRACT**

*Location specific characteristics of a road segment such as road geometry as well as surrounding road features can contribute significantly to road accident risk. A Google Maps image of a road segment provides a comprehensive visual of its complex geometry and the surrounding features. This paper proposes a novel machine learning approach using Convolutional Neural Networks (CNN) to accident risk prediction by unlocking the precise interaction of these many small road features that work in combination to contribute to a greater accident risk. The model has worldwide applicability and a very low cost/time effort to implement for a new city since Google Maps are available in most places across the globe. It also significantly contributes to existing research on accident prevention by allowing for the inclusion of highly detailed road geometry to weigh in on the prediction as well as the new location-based attributes like proximity to schools and businesses.*

## **KEYWORDS**

*Deep Learning, Convolutional Neural Networks, Maps Images, Road Accidents*

## **1. INTRODUCTION**

### **1.1. Motivation**

In 2016, 1.35 million deaths were caused by road traffic accidents worldwide. An additional 20-50 million people were injured or disabled due to a road accident. Road accidents are also the leading cause of death for children and youths ages 5-29. Developing countries, which are slowly becoming more motorized, fare much worse, with death rates three times higher than developed countries. The World Bank and the WHO have both declared that the number of road accidents is much too high in both developed and developing countries and that governments must take steps to reduce them.

Apart from the social cost of road accidents, there is also a massive economic cost. Many studies concur that road accidents can cost countries 2% of their GDP. The World Bank stated that “halving deaths and injuries due to road traffic could potentially add 22% to GDP per capita in Thailand, 15% in China, 14% in India, over 2014-2038.” It is clear that road accidents pose a significant economic barrier for developing nations. Thus, reducing road accidents and improving road safety is of paramount importance, for both developed and developing countries.

### **1.2. Previous Work**

Many of the current approaches to reducing road accidents are aimed either at making vehicles safer or at reducing human error through awareness campaigns [1] and safety training for

drivers. Other approaches focus on real time traffic flow prediction [2,3] and individual risk prediction [4]. This research shifts away from the existing work and looks at the road itself to develop a possible approach to identify accident prone road segments using Google Maps Images and a Deep Learning approach based on Convolutional Neural Networks.

Previous research has demonstrated that road characteristics can play a role in causing road accidents. [5] found out through an analysis of crash data that curves, shoulders and super elevation were more strongly related to crash rate. Though these characteristics may seem common sense, the research also discovered very specific combinations of road characteristics which contributed to crash rates (ie. lane changing crashes on larger right curves). If more of these specific combinations could be identified, then road engineers would be able to design roads more safely without these combinations of characteristics.

In addition, there has been research to suggest that location features other than road characteristics can contribute to accident risk. [6] found that community design plays a role in accident risk. They identified that big-box stores, arterial thoroughfares, and strip commercial uses were more dangerous than pedestrian scaled retail uses. This may indicate that features next to road segments (ie. strip malls, retail stores) may make those segments inherently more dangerous for drivers. Thus, a full analysis of the road segment and its surrounding features may allow us to make a more accurate assessment of a road accident risk.

There has already been some research in this field of identifying high risk crash sites. [7] carried out a study of crash sites in California to create a model to identify high risk road segments based on road characteristics. Software like SafetyAnalyst can also rank order sites based on frequency of crashes. Though these programs are useful to an extent, there are some problems with these approaches that reduce their effectiveness. One, they do not consider surrounding features like malls and restaurants which could play a role in increasing accident risk. Two, they heavily rely on data that government agencies have collected. Many local governments do not have the resources to collect this type of data and as a result, any available data is usually unreliable or incomplete. Three, the data these models rely on are linear discrete values, which ignore the complexity of road intersections and characteristics. For example, instead of observing how a curve is oriented in relation to other road characteristics, the data contains a single number indicating the angle of the curve. Finally, these models are simplistic, and ignore deep interactions between road characteristics which could influence the accident rate.

### **1.3. Proposed Solution**

This paper overcomes the above limitations by training a convolutional neural network (CNN) model using past accident data and Google Maps images of accident road segments. The CNN trains by relating accidents to features present in a Google Maps image (road characteristics, location features, for example see Figure 1). In the end the model outputs a reliable accident risk score for a Google Maps image of any location. This paper uses data from cities of Austin, Chicago and New York to train and test the proposed model.

Convolutional neural networks are a special type of neural network that focus specifically on image analysis. CNNs can detect patterns not discernible to humans in images and make predictions based off of them. In this case, a CNN can unlock deep interactions between the many complex features captured by a Google Maps image and use them to make more accurate assessments of accident risk. It is possible that a combination of proximity to a certain type of business and a certain road characteristic is much more dangerous than another such combination. The CNN should be able to tell the difference and produce an accurate risk score for both.

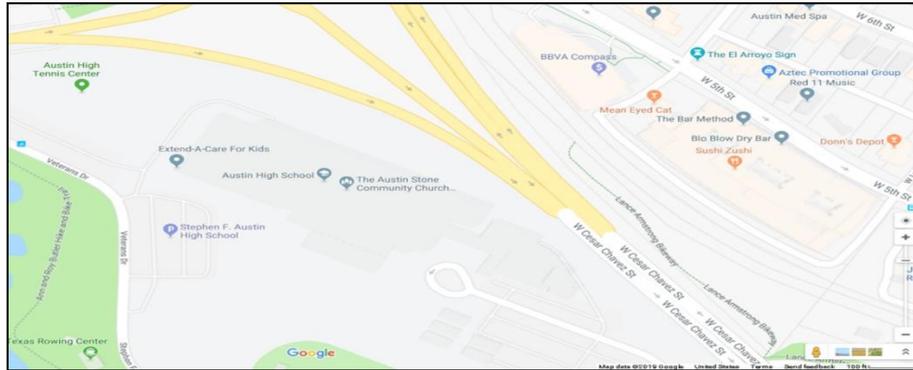


Figure 1: Google Maps Image

Further, a Google Maps image as input for this model is especially useful for the model to be universally applicable. While government agencies, especially in developing countries, may or may not collect crash data reliably, Google continuously updates their Maps information for navigation, so the data required to feed the model is present almost everywhere. It is readily accessed through the API Google provides worldwide. This makes the approach presented in this paper very cost effective and universally available.

## 2. MATERIALS, METHODS, AND PROCEDURE

This paper develops a CNN model to predict accidents using Google Maps images. Details of the model, data and procedure are included below:

### 2.1. CNN

CNNs work through passing multiple filters randomized with weights across a pixel array of an image. These filters capture different features of the image and slowly downsize the image until it produces a single numerical output. The weights in the filters are adjusted as the network is trained until the network can accurately identify the image label. In the context of this paper, the CNN learns the weights for filters to identify certain road features and location attributes (i.e. curves and icons for businesses). When the model downsizes the image, it is identifying certain combinations of these features that cause a higher risk of an accident.

### 2.2. Google Maps Images

These capture many relevant features and very detailed road geometry characteristics that can contribute to road accident risk. Figure 1 shows that Google Maps images provide a comprehensive visual of the number of intersecting roads, the angle of intersection, the sharpness of curvature, the direction of curvature, and the road type and width. If one simply took the linear data captured by government agencies about simple curvature and intersection, then they could not make as accurate a representation of the road as is represented in Google Maps. Most of this complexity is lost when a road is taken and made into linear, discrete data, but all of this complexity is retained when it is kept in the format of a Maps image.

In addition to the road characteristics, a significant advantage of using Google Maps images is the surrounding features that are displayed. Looking at Figure 1, it is clear that the image displays businesses (and through icons the nature of those businesses), schools, gas stations,

restaurants, malls, bike trails, the shape and size of nearby buildings, parking lots, parks, and natural features like forests and water bodies. Further, many other features such as speed limit and traffic flow can be inferred and learned indirectly from the image. For example, speed limit can be predicted through the road type. On Google Maps, highways are colored yellow while normal streets are colored white. One can obviously tell that a yellow road segment probably has a higher speed limit than a white colored segment. The same can be said for traffic flow. While Google Maps does not directly display it, it can be inferred through Google’s use of arrows to indicate lane direction, the width of the road, and the presence of nearby businesses (more businesses would probably mean more traffic).

### 2.3. Procedure

Steps involved in creating the CNN are as follows

1. Download relevant crash data
2. Capture Google Maps Images
3. Prepare data for feeding through model
4. Building and Training the model
5. Improving performance of the model
6. Assessing the final performance of the model and testing scalability

Accident data was downloaded from the NYC Open Data Portal for the City of New York<sup>1</sup>. The data included 658,309 crash incidents over a time span from 2014-2019. There were 58,494 unique accident locations in the data. The fact that there were more unique incidents than locations indicates that in some areas, accidents occurred more than once. Each crash recorded contained data for several attributes related to the accident, including the road type, the exact location in latitude and longitude, and whether the crash was at an intersection. A graph of the data distribution for New York is shown in Figure 2. The mean number of accidents per location was 11.25, the median was 4.5 and standard deviation was 18.86.

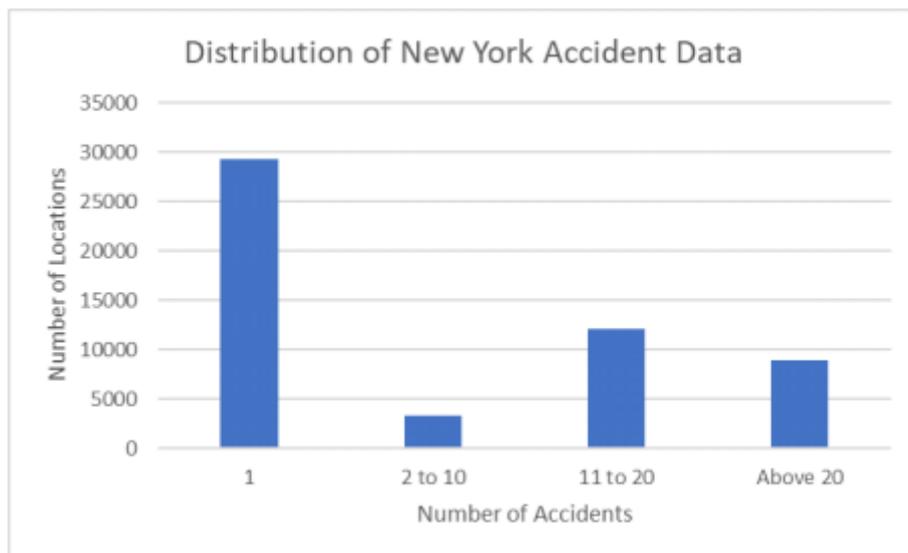


Figure 2. New York Accident Data Distribution

For each of the unique locations in the original City of New York data, a Google Maps image was required as input in the model. To obtain these images, a Python script looped through each

location in the crash data and made a call to Google Maps Static API to download the Google Maps image of the location using the latitude and longitude coordinates present in the data. The script also specified the zoom level on the maps, the scale of the image, and the type of Google Maps image.

The type of image that was settled upon was “roadmap” (others like “satellite” and “terrain” did not display surrounding features). The scale chosen was 50x50 pixels. There was a trade-off when specifying the zoom level of the image. If the image was too zoomed in, then complexity of surrounding features was ignored by the model, while if it was too zoomed out, the CNN lost focus on the actual accident site characteristics. This conclusion was verified after testing the CNN on some sample data with different zoom levels. Zoom level 18 was selected because the CNN achieved the highest accuracy with it. From Figure 3 it is clear that though some surrounding complexity was lost because of zooming in, major features in the image are still present, such as the intersection, curvature of the road, traffic flow, surrounding streets, and a couple businesses.



Figure 3. Google Maps Image with Zoom level 18

Google Maps images were present for each crash location, but there were no negatively labeled images to train the network. That is to say that there were no Google Maps images for locations which had never had an accident from 2014 to 2019. Feeding only positively labeled data will cause the CNN to learn to predict a high accident risk score for every image passed through it. One solution to this is finding random road segments across the city that had no accidents whatsoever. However, there was a simpler solution with the data already available. Locations with only one accident can be labeled as a negative, while locations with more than one accident can be labeled as positive. The reasoning behind this labeling scheme is that locations with only one accident are most likely not inherently dangerous locations. Any accident caused at those locations was probably due to human error or some other non-location factor. However, a location where accidents repeatedly occur indicates that there is some factor in the location itself which makes it more accident prone. For each location, an aggregate sum of all the accidents that occurred in the location was calculated. Then the locations were binned – locations with more than one accident were labeled as 1, while locations with only one accident were labeled as 0.

To perform holdout cross-validation, data was split into training (93%), validation (3.5%), and testing (3.5%) buckets. The training data was used to train the network. The validation data was used to assess the effectiveness of training and adjust the model architecture and hyperparameters to improve accuracy. The testing data was used to perform a final test in order to assess the model’s performance in the real world.

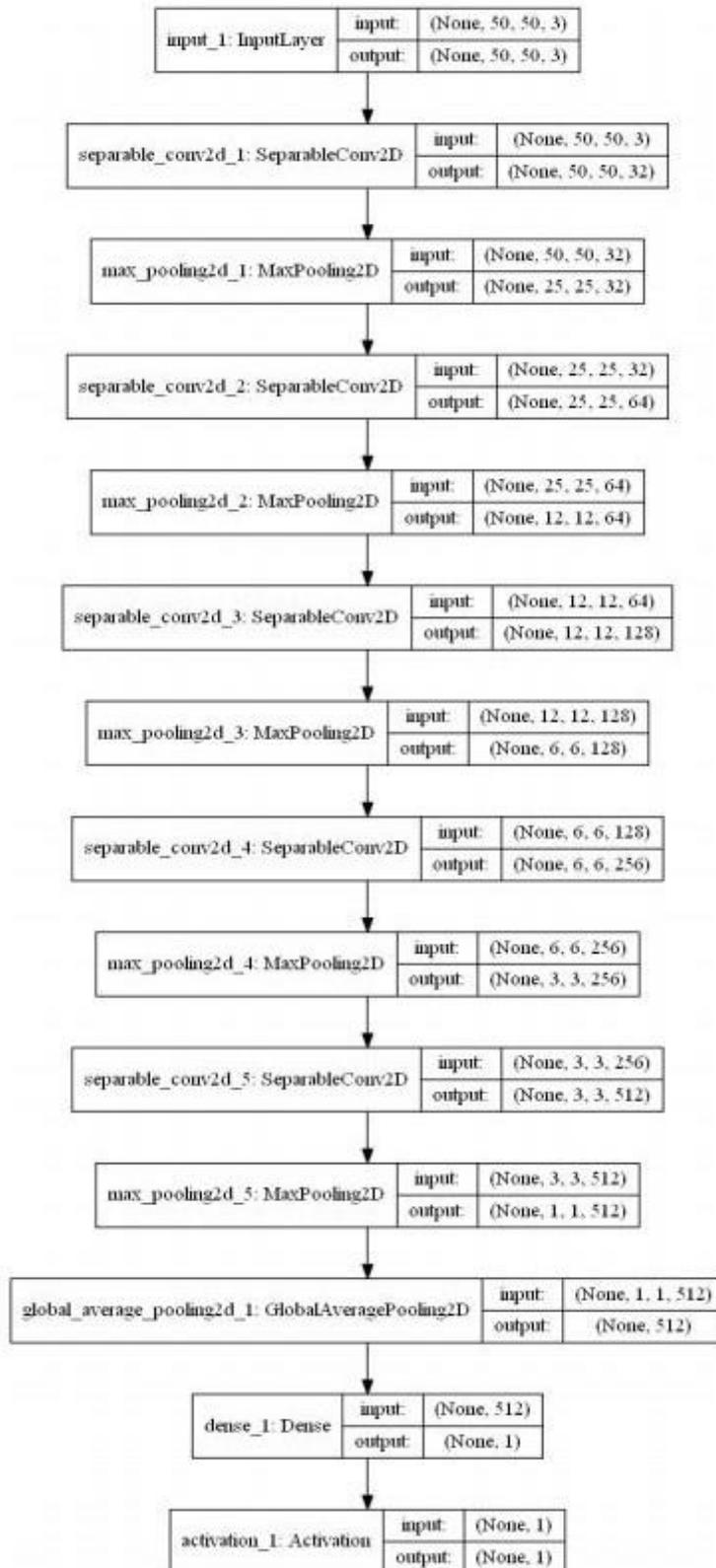


Figure 4: Architecture of CNN

To build and train the CNN, the Keras library in python was used with a TensorFlow backend. The network was trained on an AWS EC2 p3.2xlarge instance. The architecture of the network is laid out in Figure 4. The input is a Google Maps image of a location, and the output is a probability prediction of the label (specifically the probability of the location having repeat accidents from the years 2014-2019). In total the network had five convolutional layers, and each of the convolutional layers had a ReLU activation function and underwent batch normalization to allow for quick convergence. The final activation function was a sigmoid to ensure the network outputted a probability between 0 and 1. The loss function was binary cross-entropy, and the model was trained for 60 epochs using a batch size of 32.

After the CNN was completed, attempts to improve its accuracy were made. The image zooming and scaling was adjusted and the hyperparameters and architecture were optimized. Class imbalance was an additional issue. The data fed into the model was imbalanced as there were fewer data points labelled as one (a repeat accident location) as compared to data points labeled as zero (single accident location). This led to the network simply taking a naïve approach and always predicting zero. This problem was solved by weighting the neural network loss function to penalize guessing zero higher. This way the neural network was able to treat both classes equally.

### 3. RESULTS

#### 3.1. Results for Initial New York Model

All results and the performance scores were based on the test data set apart in the cross-validation process, so the model was evaluated in real world circumstances on data it had never seen before. All metrics and graphs were obtained using the scikit-learn library. The test data fed into the model contained an equal representation of each class. Finally, the accuracy scores were obtained at the optimum threshold (considering other metrics such as precision and recall).

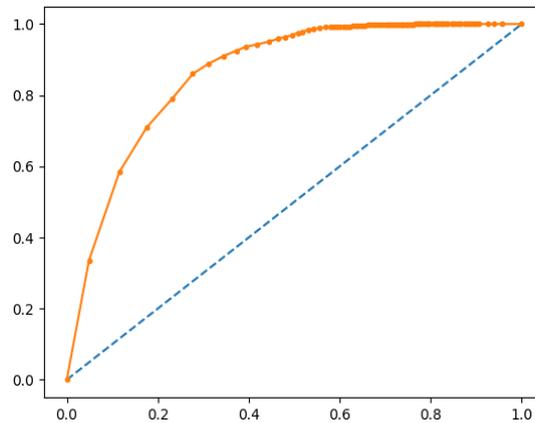


Figure 5. ROC Curve for the City of New York

The ROC curve for New York is shown in Figure 5. The AUROC was 0.93 and the accuracy was 85%. The dotted line in the middle represents a completely naïve classifier. Table 1 is the confusion matrix for the New York model and Table 2 contains precision, recall, and f1 scores selected for an optimum threshold. The support for each class is close to equal and is also large, showing that these scores reflect the true performance of the model.

Table 1. New York Model Confusion Matrix

Class	Actual 0	Actual 1
Predicted 0	725	248
Predicted 1	57	971

Table 2. New York Model Class Based Scores

	Precision	Recall	F1-Score	Support (number of examples)
Class 0	0.93	0.75	0.83	973
Class 1	0.80	0.94	0.86	1028

Table 3. Sample of Risky Locations in New York

Latitude, Longitude	Street Address	Accident Risk Score Generated By Model	Actual Number of accidents in last 6 years
40.769115, -73.98856	W 57 St/10 Av New York, NY 10019	1.0	88
40.76211, -73.96016	1150 1st Avenue #501, New York, NY 10065	1.0	65
40.758972, -73.91901	Broadway/Steinway St Queens, NY 11103	1.0	99
40.8934115, -73.8573621	4201 White Plains Rd The Bronx, NY 10466	1.0	65

A sample of risky locations in New York is shown in Table 3. All of these locations had above 60 accidents from 2014-2019, and the accident risk score of 1.0 predicted by the model indicates that the model is 100% certain that the locations are risky.

### 3.2. Robustness Analysis

To ensure that this performance was not just limited to something unique about New York, the model was retrained using data from Austin and Chicago and tested in those cities. Results for Austin are based on 66000 crash locations from Texas DOT crash data (2011-2018)<sup>2</sup> (shown in Figure 6). The accuracy achieved was 86% and the AUROC was 0.86. Results for Chicago are based on 40000 crash locations from CPD crash data (2013-2019)<sup>3</sup> as shown in Figure 6. The accuracy achieved 70% and the AUROC was 0.75. The high AUROC and accuracy across multiple tested cities suggest that the model has excellent reliability, scalability, and rank

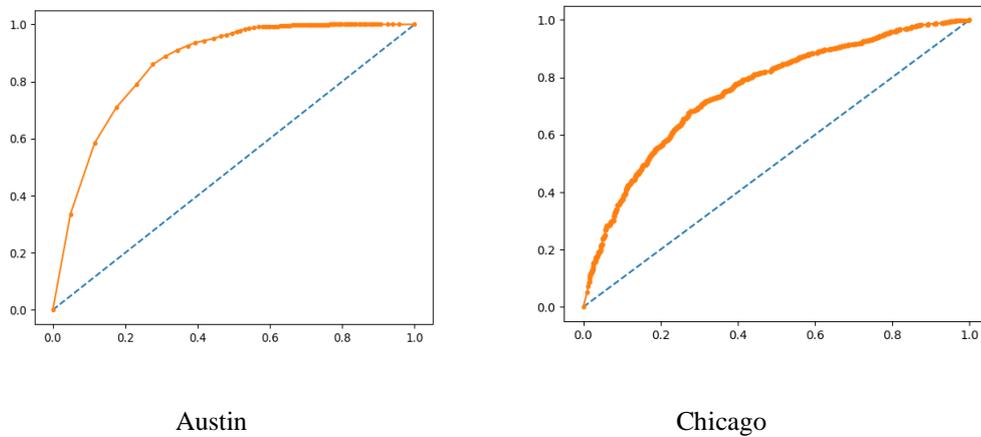


Figure 6. ROC Curves for Austin and Chicago

## 4. DISCUSSION AND CONCLUSIONS

A CNN model was developed to predict accident prone road segments using crash data from different cities and Google map images. The model achieved the design criteria stated earlier. It captures road and surrounding characteristics and their interactions and is cost effective and easy to implement.

### 4.1. Contribution

The contribution to existing research on accident prevention is significant. This paper proposed a novel machine learning approach using Google Maps images. The use of Google Maps images allows for the inclusion of highly complex road geometry which is not currently captured by models. In addition, the model allows for novel factors to be included in accident prevention, like proximity to schools, businesses and restaurants.

There are multiple potential applications of the proposed model. It allows for optimal distribution of limited city resources. Since cities can rank order locations from most dangerous to least accurately, they can prioritize placing warning signs, police officers and speed monitors at more dangerous locations. [8] explain that placing police officers in the right locations can significantly reduce fatality rates in accidents. Self-driving cars and trucks can auto-calculate risk scores for the current road they are on using the model and alter their driving behavior (perhaps drive slower when at high-risk sites). Cities can also now identify very high-risk road segments using the model and then divert funds to redesign them.

A prototype of a real time accident prediction model was created by the author using weather data (rain, wind, visibility, temperature)<sup>4</sup>. Once the output of the CNN was fed into this model, the accuracy jumped from 37% to 72%. This shows that the output of this model can be fed into other accident prediction models to improve their performance. Real time accident prediction models can help place emergency resources at the right locations before accidents even happen.

Finally, the author created a prototype consumer android app for safe driving. It calculates a risk score for all feasible routes to a destination using the model and maps the safest route. It could

International Journal of Artificial Intelligence & Applications (IJAA) Vol.10, No.6, November 2019  
be enhanced to warn drivers on risky segments so distracted driving in risky segments is reduced. An app like this could be especially helpful for senior citizens, teenage drivers, and people who are driving in unfamiliar locations.

Proposed solution is also very cost effective. The cost of obtaining Google Maps images would be around \$200 for each city (depending on the city size). The economic benefits gained by preventing crashes far outweighs the negligible cost. Cities can choose to retrain the network for themselves (in which case they would need data that gives crash locations over a certain time period) or opt to just use a network that works in another city. Once the model scripts were created, it took very little time to implement the model for Austin and Chicago. The model also does not require any manual updates. It can be run once every year to rank order locations according to the newest Google Maps images. Google will continue to update Maps, so any changes in road networks will be accounted for.

Overall, the above use case scenarios are applicable anywhere in the world Google Maps is available including the developing countries.

## **4.2. Limitations and Future Work**

There are a few limitations when using this model. The first is that the model does not consider driver and vehicle characteristics. Even on very safe roads it is possible for a drunk driver to cause an accident. In addition, the Google Maps view is limited to a two-dimensional bird's eye view of the road segment. There is research to suggest that three dimensional features on roads can contribute to accident, such as specific roadside vegetation and road shoulders. The model cannot take this three-dimensional data into account. Another limitation is that the model does not train based on accident severity. It considers all accidents to be the same regardless of the number of fatalities involved or the damage caused.

This research can be expanded in multiple ways. Real time traffic data could be included in the model. The model could also be trained to predict fatalities and the severity of accidents. Three-dimensional data can be considered by using programs like Google Street View. Another extension would be to build an end-to-end system where city officials can gather data, train and predict accidents with a few clicks.

## **REFERENCES**

- [1] Hoekstra, Tamara, and Fred Wegman. "Improving the Effectiveness of Road Safety Campaigns: Current and New Practices." *IATSS Research* 34, no. 2 (2011): 80–86. <https://doi.org/10.1016/j.iatssr.2011.01.003>.
- [2] Polson, Nicholas G., and Vadim O. Sokolov. "Deep learning for short-term traffic flow prediction." *Transportation Research Part C: Emerging Technologies* 79 (2017): 1-17.
- [3] Zhang, Zhenhua, Qing He, Jing Gao, and Ming Ni. "A deep learning approach for detecting traffic accidents from social media data." *Transportation research part C: emerging technologies* 86 (2018): 580-596.
- [4] Chen, Quanjun, Xuan Song, Harutoshi Yamada, and Ryosuke Shibasaki. "Learning deep representation from big and heterogeneous data for traffic accident inference." In *Thirtieth AAAI Conference on Artificial Intelligence*. 2016.
- [5] Othman, S., Thomson, R., & Lannér, G. (2009, October). Identifying critical road geometry parameters affecting crash rate and crash type. In *Annals of Advances in Automotive*

International Journal of Artificial Intelligence & Applications (IJAA) Vol.10, No.6, November 2019  
Medicine/Annual Scientific Conference (Vol. 53, p.155). Association for the Advancement of  
Automotive Medicine.

- [6] Dumbaugh, Eric, Yi Zhang, and Wenhao Li. Community design and the incidence of crashes involving pedestrians and motorists aged 75 and older. No. UTCM 11-03-67. Texas Transportation Institute. University Transportation Center for Mobility, 2012.
- [7] Geyer, Judy, Elena Lankina, Ching-Yao Chan, David Ragland, Trinh Pham, and Ashkan Sharafsaleh. "Methods for Identifying High Collision Concentration Locations for Potential Safety Improvements." CALIFORNIA PARTNERS FOR ADVANCED TRANSIT AND HIGHWAYS, December 2008.
- [8] Rezapour, Mahdi, Shaun S. Wulff, and Khaled Ksaibati. "Effectiveness of Enforcement Resources in the Highway Patrol in Reducing Fatality Rates." IATSS Research 42, no. 4 (2018): 259–64. <https://doi.org/10.1016/j.iatssr.2018.04.001>.

## **AUTHORS**

Aarya Agarwal is a student at Westwood High school, Austin and he is currently pursuing an IB diploma. He has completed several machine learning projects and has submitted several of these projects to prestigious science competitions. His achievements include 1st place at the Texas State Science Fair and Austin Regional Science Fair, as well as 2nd place at the Texas Junior Academy of Sciences for the category of Computer Science/Math

