Combining Satellite Imagery and Open Data to Map Road Safety

Alameen Najjar, Shun'ichi Kaneko, Yoshikazu Miyanaga
Graduate School of Information Science and Technology, Hokkaido University, Japan
najjar@hce.ist.hokudai.ac.jp, {kaneko, miya}@ist.hokudai.ac.jp

Abstract

Improving road safety is critical for the sustainable development of cities. A road safety map is a powerful tool that can help prevent future traffic accidents. However, accurate mapping requires accurate data collection, which is both expensive and labor intensive. Satellite imagery is increasingly becoming abundant, higher in resolution and affordable. Given the recent successes deep learning has achieved in the visual recognition field, we are interested in investigating whether it is possible to use deep learning to accurately predict road safety directly from raw satellite imagery. To this end, we propose a deep learning-based mapping approach that leverages *open data* to learn from raw satellite imagery robust deep models able to predict accurate city-scale road safety maps at an affordable cost. To empirically validate the proposed approach, we trained a deep model on satellite images obtained from over 647 thousand traffic-accident reports collected over a period of four years by the New York city Police Department. The best model predicted road safety from raw satellite imagery with an accuracy of 78%. We also used the New York city model to predict for the city of Denver a city-scale map indicating road safety in three levels. Compared to a map made from three years' worth of data collected by the Denver city Police Department, the map predicted from raw satellite imagery has an accuracy of 73%.

Introduction

In a recent report (World Health Organization 2015), the World Health Organization (WHO) has estimated that road traffic accidents are the number one cause of mortality among people aged between 15 and 29 years, killing more than 1.2 million people worldwide each year. Most of road deaths are in low- and middle-income countries where infrastructure development and policy change have not kept pace with the increase in motorization accompanying the rapid economic growth. Poor road safety causes developing countries an approximate economic loss of 3% of GDP making it not only a major public health issue but also a developmental one. In the face of this grim reality, the world's nations have recently adopted the 2030 agenda of Sustainable Development which sets the goal of cutting road injuries and deaths in half by 2020 (United Nations 2015).
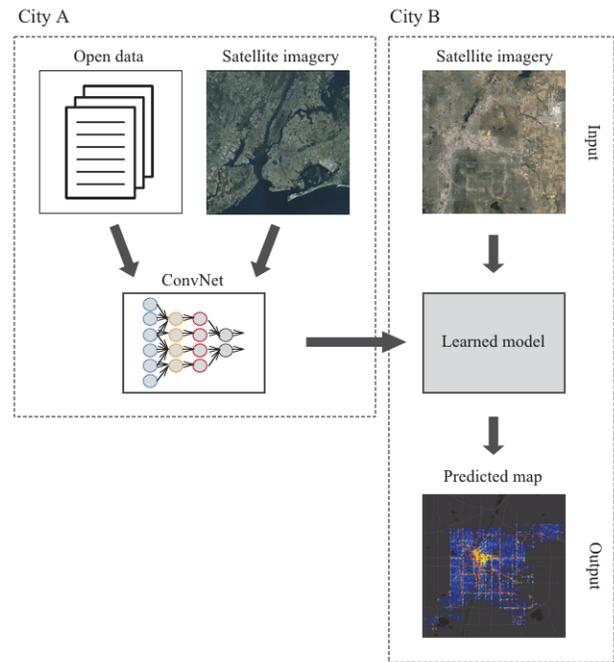
Figure 1: City-scale road safety mapping from satellite imagery via deep learning.

Mapping traffic accidents is an established practice used to gain insights on where and what interventions are needed to improve road safety (Miaou, Song, and Mallick 2003). A map made from manually collected reports of previous accidents visualizes where within the city road safety suffers. Maintaining and improving infrastructure around these spots helps prevent future traffic accidents.

However, accurate mapping requires accurate data collection, which is both expensive and labor intensive. While high-income countries are flooded with data, most low- and middle-income countries suffer from a data drought (Leidig, Teeuw, and Gibson 2016). The reason is that accurate data collection requires transparency, technology, skills and other resources extremely scarce in these parts of the world. Therefore, an automatic approach to road safety mapping that does not require data collection is highly needed.

(a)	(b)	(c)
(d)	(e)	(f)

Figure 2: Satellite images of six different locations in New York city. Between March 2012 and March 2016, locations in the upper row had over 100 traffic accidents each. Those in the bottom row had only one accident each. What is interesting is the striking visual similarity among images of the same row. Notice how images of locations of similar road safety level have similar (1) setting (highway/intersection vs. residential), (2) dominant color (gray vs. green), and (3) objects (zebra lines and vehicles vs. trees and rooftops). This example illustrates that visual features captured in satellite imagery can be used as a proxy indicator of road safety. Therefore, based mainly on this observation we are motivated to investigate predicting road safety directly from raw satellite imagery.

Recent advances in imaging and space technology have increasingly made satellite imagery abundant, higher in resolution and affordable (Dash and Ogutu 2016). The bird's eye/aerial viewpoint of satellite imagery makes it a rich medium of visual cues relevant to environmental, social, and economic aspects of urban development. Given the recent successes deep learning (LeCun, Bengio, and Hinton 2015) has achieved in the visual recognition field (Razavian et al. 2014; Oquab et al. 2014; Donahue et al. 2014), we are interested in investigating the use of deep learning to predict road safety from raw satellite imagery (See Figure 1).

Although satellite imagery captures a wealth of visual features relevant to road safety (See Figure 2 for an illustrated example), it is a highly *unstructured* form of data. Given that learning robust models using the recently proven successful deep networks (Krizhevsky, Sutskever, and Hinton 2012) requires large numbers of training examples, it is difficult to use deep learning to extract useful insights on road safety directly from raw satellite images.

We overcome this problem by leveraging an abundant yet highly accurate source of data known as *open data* (Dietrich et al. 2009). The idea is to mine available large-scale datasets of traffic-accident reports for high-quality labeled satellite images. These datasets are official (i.e., collected by governments) and made publicly accessible online. To the best of our knowledge, we are the first to adopt such strategy in collecting and labeling satellite images.

In this paper, we propose a deep learning-based mapping approach that uses Convolutional Neural Networks (ConvNets) (LeCun et al. 1989; 1998) and leverages open data to learn from raw satellite imagery robust deep models able to predict accurate city-scale road safety maps. A deep model learned from open data collected in one city is used to generate for another city a road safety map predicted from its satellite imagery (See Figure 1). Since data collection is not required, our approach offers a mapping solution that is highly affordable.

To empirically validate the proposed approach, we trained a deep model on satellite images obtained from over 647 thousand traffic-accident reports collected over a period of four years by the New York city Police Department (NYPD). The best model predicted road safety from raw satellite imagery with an accuracy of 78%. To test its reusability, we used the New York city model to predict for the city of Denver a city-scale map indicating road safety in three levels. Compared to a map made from three years' worth of data collected by the Denver city Police Department, our map has an accuracy of 73%.

To the best of our knowledge, this paper is the first to attempt predicting road safety directly from raw satellite imagery. Contributions made in this paper are summarized as follows:

1. Proposing a deep learning-based mapping approach for predicting city-scale road safety maps from raw satellite imagery.

2. Proposing the idea of obtaining labeled satellite images from open data.

3. Making publicly available a deep model for predicting road safety learned from over 647 thousand traffic-accident reports collected by the NYPD.

4. Predicting for the city of Denver a city-scale road safety map with an accuracy of 73%.

The remainder of this paper is organized as follows. Previous works on machine learning-based road safety mapping are briefly reviewed in section 2. Data is introduced in section 3. Our approach is explained in section 4. Empirical validation using real data is presented in section 5. Finally, the paper is summarized and concluded in section 6.

Previous Works

In this section, we briefly review previous works on city-scale road safety mapping using machine learning, and compare them to ours.

To the best of our knowledge, (Chen et al. 2016) is the only work that uses machine learning to predict city-scale road safety maps. In this work, a deep model is learned from traffic-accident reports and human mobility data (i.e., GPS data) collected from 1.6 million smartphone users over a period of seven months. The learned model is then used to predict from real-time GPS data a map for the city of Tokyo indicating road safety in three different levels.

(Chen et al. 2016) is similar to our work in the fact that it uses patterns recognized in an abundant and unstructured source of data as a proxy indicator of road safety. While

ID	Date	Time	Latitude	Longitude	Vehicle 1	Vehicle 2
1	3/12/2016	10:30	40.5426730	-74.1647651	Station wagon	Van
2	3/12/2016	12:15	40.5970318	-74.0933639	Station wagon	Unknown
3	8/31/2015	09:40	40.6338578	-74.1259566	Passenger vehicle	Bus
4	8/29/2015	07:08	40.6134776	-74.0979215	Unknown	Other
5	8/19/2015	08:30	40.6311355	-74.1279294	Passenger vehicle	Bicycle

Table 1: Examples of NIBRS-style traffic accident reports.

Chen et al. use real-time GPS data, we use satellite imagery as our abundant source of data. However, the core difference between the two works is the application domain each is intended for. While Chen et al. are interested in generating user-oriented maps intended for real-time use, we are interested in generating maps that help inform city-planning decision-making and policy and eventually improve road safety for cities where proper data collection is unaffordable.

It is worth mentioning that for the application we are interested in, using satellite imagery rather than GPS data is more practical since: (1) satellite images are more ubiquitous and freely available online (e.g., from Google Maps), and (2) smartphones in low- and middle-income countries (which this research is targeting) are not as widely used as in high-income countries, i.e., GPS data in low- and middle-income countries do not provide a reliable indicator of road safety at a city scale.

We are aware of other works, such as (Anderson 2009; Bíl, Andrášik, and Janoška 2013; Xie and Yan 2013; Han et al. 2015), which mainly focus on the detection and analysis of traffic accident-prone areas (also known as, traffic accident hotspots) rather than the prediction of road safety level at a city scale. Therefore, and given the above, we believe that our work is the first to attempt using machine learning to predict city-scale road safety maps directly from raw satellite imagery.

Data

In this section, we describe the data we used to train, verify, and test our deep models.

All models in this paper were trained, verified and tested on satellite images obtained from official traffic-accident reports collected by police departments in the United States. These reports are released as *open data* which is defined as data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share-alike (Dietrich et al. 2009).

Reports follow the National Incident Based Reporting System (NIBRS) (Maxfield 1999) that describes individual accidents using several attributes, such as time, date, (approximate) geographic location of the accident, and types of vehicle involved. See Table 1 for an example.

We used data collected in two US cities (New York and Denver), and it is summarized as follows:

- 647,868 traffic-accident reports collected by the NYPD over the period between March 2012 and March 2016[1].

- 110,870 traffic-accident reports collected by the Denver city police department over the period between July 2013 and July 2016.

Given that these reports were collected manually by human experts (e.g., police officers), we assume that models learned from this data are highly accurate. Using the geographic location information, we represented individual reports with satellite images. We used these images to train, verify, and test our deep models.

Deep Models for Road Safety Mapping

The ability of mapping road safety implies the ability of predicting the occurrence of traffic accidents. However, traffic accidents occur due to complex reasons which cannot be pinned down to a single one. Therefore, it is extremely challenging to predict traffic accidents.

In this work, we base our definition of road safety on the concept of accident hotspots. In other words, we predict the level of road safety based on both frequency and severity of previous accidents occurred within a limited geographic area. Assuming that satellite imagery captures a wealth of visual information that can be used as a proxy indicator of road safety, we aim to model and understand what a satellite map of a city tells about road safety.

In the following, we describe our approach. First, we explain how we obtain labeled satellite images from raw open data. Then, we describe how we use ConvNets to train deep models for predicting road safety from satellite imagery.

Satellite Images from Open Data

Before learning a robust deep model able to predict road safety from raw satellite images, first we need to collect a set of training images accurately labeled with road safety, and of a scale large enough to train a deep network. To this end, we propose to obtain our images from an abundant yet accurate source of data known as open data. This procedure is explained in the following:

Location information discretization: using a square grid, we divided the map of New York city into square regions (r) of 30m×30m each which is a proper area for traffic accident analysis. Then given their location information, the 647,868 traffic accidents documented by the NYPD were assigned to different regions. Finally, each region is assigned a safety score (S_r) given as the sum of severity levels of accidents occurred within its boundaries, such that:

$$S_r = \sum_{i=1}^{n} a_{i,r}, \tag{1}$$

where $a_{i,r}$ is the severity level of the i-th accident within region r, and n is the total number of accidents. $a_{i,r} = 1$ whenever severity level is unknown.

Binning: to obtain three safety levels (high, neutral, and low), we clustered the obtained safety scores by frequency around three bins using the k-means algorithm (MacQueen and others 1967), such that:

$$\arg\min_T \sum_{i=1}^{k} \sum_{x \in T_i} \|x - \mu_i\|^2, \qquad (2)$$

where μ_i is the mean of the points in T_i, $k = 3$ is the number of bins, and x is the frequency of individual scores. We have experimented with other clustering algorithms, such as Gaussian Mixture Models (GMM) and Jenks natural breaks optimization (Jenks 1967). However, we found that the simple k-means gives the best results.

Resampling: the obtained labeled set of regions is highly imbalanced. In fact over 59% of the regions are labeled as highly safe. Therefore and in order to avoid learning a biased model, we resampled our data via downsampling majority classes so that the three classes are balanced out.

Finally, we represented each of the regions with a satellite image centered around the GPS coordinates of its center. These images form a dataset on which our models are trained, verified and tested.

Safety Prediction using ConvNets

We begin by introducing ConvNets, then we explain how our models are learned.

Convolutional Neural Networks A ConvNet is a biology-inspired feedforward neural network that is designed to process data that come in multiple arrays, such as RGB color images. Similar to other deep learning approaches, ConvNets automatically learn from data hierarchical representations that capture patterns and statistics at multiple levels of abstraction.

Having their roots in the early neocognitron (Fukushima and Miyake 1982), ConvNets have been used in several applications since the early 1990s such as in (LeCun et al. 1998). Later in the 2000s, ConvNets proved highly successful in several vision tasks where training examples are abundant. However, not until 2012 when trained on over a million images, ConvNets achieved a ground-breaking performance in generic object recognition. This success has since revolutionized in the visual recognition field, with ConvNets dominating most of the vision tasks nowadays (LeCun, Bengio, and Hinton 2015).

A ConvNet takes a raw RGB image as an input and produces a class prediction as an output. Natural images are compositional hierarchies, in which lower level features combine to form higher level ones. ConvNets were designed to exploit this property. A typical ConvNet consists of a stack of convolutional layers followed by fully-connected layers ordered such that the output of one layer is the input of the next. A typical convolutional layer convolves a three-dimensional input tensor with a tensor of weights (filter maps). The weighted sum of the convolution is then passed through a nonlinearity function such as a Rectified Linear Unit (ReLU). The result is then passed through pooling operators (usually, max operator) to reduce the dimensionality of the representation and make it invariant to small perturbations. On the other hand, a fully-connected layer reduces the multidimensional input into a one-dimensional vector that is fed to a final classifier.

A ConvNet is trained end-to-end in a supervised fashion using stochastic gradient descent (SGD) and backpropagation.

Model Learning To train our models, we adopted transfer learning in which pre-learned knowledge is transferred from a source to a target problem. In our case, source and target problems are generic object/scene recognition and road safety prediction, respectively. And the transferred knowledge is a set of low-level visual features such as edges and corners. In the deep learning community, this way of training is known as finetuning and it has been proven highly successful in augmenting learning when training data is limited (Karayev et al. 2013; Branson et al. 2014).

To finetune a pre-trained model, we first replaced the classification layer with a three-class output layer representing the three safety levels. Weights of the newly added layer are initialized randomly, and the entire network is trained jointly using small learning rates.

Experiments

In this section, we empirically validate the proposed approach by evaluating the accuracy of a road safety map predicted for the city of Denver using a deep model learned from data collected in New York city.

Datasets

In our experiments, we used two different satellite imagery datasets collected as explained in the previous section:

NYC: It consists of 14,000 satellite images obtained from official traffic-accident reports collected by the NYPD. Individual images are labeled with one of three safety labels.

Denver: It consists of 21,406 satellite images obtained from official traffic-accident reports collected by the Denver city Police Department. Individual images are labeled with one of three safety labels.

Implementation details

Favoring the reproducibility of the results, below we explain how experiments were implemented:

Satellite imagery: We used Google Static Maps API[2] to crawl all satellite images used in our experiments. Individual images have a spatial resolution of $256{\times}256$ pixels each and crawled at three different zoom levels (18, 19, and 20).

Architecture: All ConvNets used in this experiments follow the AlexNet architecture (Krizhevsky, Sutskever, and Hinton 2012). We are aware of the other successful deep architectures, such as (Simonyan and Zisserman 2014; Szegedy et al. 2015). However, we used AlexNet since it is both simple and considered a landmark architecture.

---

[2]https://developers.google.com/maps

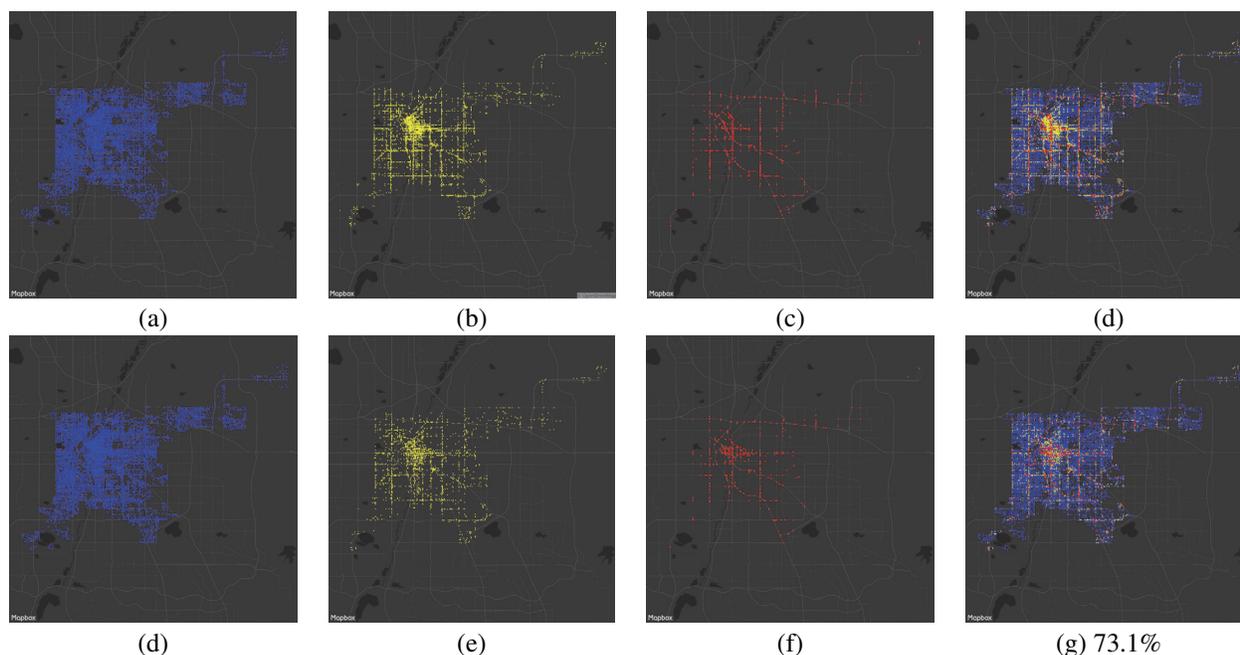| | | | |
| (a) | (b) | (c) | (d) |
| (d) | (e) | (f) | (g) 73.1% |

Figure 3: City-scale map of Denver city indicating road safety in three different levels (high: blue, neutral: yellow, and low: red). Upper row is a map made from data collected by Denver city Police Department between July 2013 and July 2016. Bottom row is a map predicted from raw satellite imagery using our approach. First three columns (left to right) represent the three safety levels mapped individually. The fourth column represents all safety levels mapped together. This figure is best viewed in digital format.

	x18	x19	x20
ImageNet	0.740	0.766	0.739
Places205	0.755	0.775	0.745
ImageNet + Places205	0.778	**0.782**	0.771

Table 2: Average prediction accuracy obtained using different models pre-trained on three different large-scale datasets and finetuned on satellite images captured at three different zoom levels. The best performing model is the one pre-trained on both ImageNet and Places205 and finetuned on satellite images captured at zoom level 19. All results are cross-validated on three random data splits.

	Denver	New York
Population (million)	0.664	8.491
Area (km^2)	396.27	783.84
Population density (people per km^2)	1675	10833
Average traffic delay (minutes per person per day)	7.4	9.7

Table 3: Comparing Denver city to New York city in terms of area, population and traffic delay.

Training: Instead of training from scratch, our models were pre-trained on a generic large-scale image dataset first. Three pre-training datasets were considered: (1) ImageNet (Deng et al. 2009), (2) Places205 (Zhou et al. 2014), and (3) both ImageNet and Places205 datasets combined. Finally, training was conducted using Caffe deep learning framework (Jia et al. 2014) running on a single Nvidia GeForce TITAN X GPU.

Evaluation: To evaluate the learned models, we calculated the average prediction accuracy cross-validated on three random 5%/95% data splits. Reported results are obtained after 60,000 training iterations.

Predicting Road Safety in New York City

The purpose of this experiment is twofold: (1) to investigate whether or not our assumption that visual features captured in satellite imagery can be effectively used as a proxy indicator of road safety. And (2) to evaluate the performance of state-of-the-art ConvNets in learning deep models able to predict road safety from raw satellite images.

We have finetuned a ConvNet on images of the *NYC* dataset. Table 2 shows the average prediction accuracy obtained using nine models obtained considering three pre-training scenarios using satellite images captured at the three zoom levels.

Spanning a range between 73.9% and 78.2%, the best performing model is the one obtained through finetuning a pre-trained model on both ImageNet and Places205 datasets using satellite images captured at zoom level 19.

From Table 2, one can make the following observations:

1. For all zoom levels, models pre-trained on both ImageNet and Places205 achieves the best, followed by models pre-trained on Places205, and finally models pre-trained on ImageNet. This is expected since satellite images have bird's eye/aerial viewpoint which makes them closer in composition to scene images of Places 205 rather than the object-centric images of ImageNet.

2. For all pre-training scenarios, finetuning using satellite images captured at zoom level 19 results in the best performance.

Results obtained in this experiment confirm our assumption that visual features captured in satellite imagery can be efficiently used as a proxy indicator of road safety. Moreover, state-of-the-art ConvNets are able to learn robust models that can predict road safety from raw satellite images.

City-Scale Mapping of Denver City

The purpose of this experiment is to evaluate the performance of a deep model learned from data collected in New York city in predicting for the city of Denver a city-scale map indicating road safety in three levels from raw satellite images only.

To this end, we used the best performing model learned from New York city to predict safety labels of the 21,406 images of the *Denver* dataset. Figure 3 shows a city-scale road safety map for the city of Denver. The upper row is a map made from 110,870 traffic-accident reports collected by the Denver police department over the period between July 2013 and July 2016. The bottom row shows a map predicted completely from raw satellite images. The first three columns (left to right) illustrate the three safety levels (high: blue, neutral: yellow, and low: red) mapped individually. The fourth column illustrates all safety levels mapped together. Compared to the official map (upper row), the predicted map (bottom row) has an accuracy of 73.1%.

Denver city and New York city are quite different from each other in terms of the level of development, area, population, traffic, etc. (See Table 3 for a brief comparison). Thus, demonstrating that a model learned from New York city data can effectively predict road safety in Denver city proves that models are practically reusable. Moreover, in order to quantify the accuracy of the predicted map, we had to choose a city that has its official traffic-accident reports publicly accessible so that we can compare our results to a ground truth. Therefore, for the previous reasons we chose Denver city to map in this experiment.

Results obtained in this experiment confirm that deep models learned from road safety data collected in a large city can be reused to predict road safety in smaller cities with less resources.

Summary & Conclusions

In this paper, we have investigated the use of deep learning to predict city-scale road safety maps directly from satellite imagery. We have proposed a mapping approach that uses state-of-the-art Convolutional Neural Networks (ConvNets) and leverages open data to learn robust deep models able to predict road safety from raw satellite imagery.

To empirically validate the proposed approach, we trained a deep model on satellite images obtained from over 647 thousand traffic-accident reports collected over a period of four years by the New York city Police Department. The best model predicted road safety from raw satellite imagery with an accuracy of 78%. We also used the New York city model to predict for the city of Denver a city-scale map indicating road safety in three levels. Compared to a map made from three years' worth of data collected by the Denver city Police Department, the map predicted from raw satellite imagery has an accuracy of 73%.

The obtained results confirm: (1) our assumption that visual features contained in satellite imagery can be effectively used as a proxy indicator of road safety. (2) State-of-the-art ConvNets can learn robust models for road safety prediction from satellite imagery. (3) Deep models learned from road safety data collected in a large city can be reused to predict road safety in smaller cities with less resources.

Although providing a practical and affordable approach for road safety mapping where proper data collection is not affordable (e.g., developing countries), our study suffers from several limitations. First, our models were trained on traffic-accident reports that do not indicate the severity level of individual accidents. Therefore, we used accident frequency only as safety labels. We believe that training models on more elaborate data will result in better performance. Second, our models predict road safety without taking time and date into consideration. In other words, our maps do not differentiate between day and night or summer and winter. Third, although we proved our method effective in predicting road safety in Denver city using models trained on data collected in New York city, we have not considered a more extreme case in which both cities are located in two different continents (e.g., New York city and Nairobi city) where architecture, city planning, traffic regulations (among others) differ extremely. These limitations among others are to be addressed in future work.

References

Anderson, T. K. 2009. Kernel density estimation and k-means clustering to profile road accident hotspots. *Accident Analysis & Prevention* 41(3):359–364.

Bíl, M.; Andrášik, R.; and Janoška, Z. 2013. Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. *Accident Analysis & Prevention* 55:265–273.

Branson, S.; Van Horn, G.; Belongie, S.; and Perona, P. 2014. Bird species categorization using pose normalized deep convolutional nets. *arXiv preprint arXiv:1406.2952*.

Chen, Q.; Song, X.; Yamada, H.; and Shibasaki, R. 2016. Learning deep representation from big and heterogeneous data for traffic accident inference. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Dash, J., and Ogutu, B. O. 2016. Recent advances in space-borne optical remote sensing systems for monitoring global terrestrial ecosystems. *Progress in Physical Geography* 40(2):322–351.

Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 248–255. IEEE.

Dietrich, D.; Gray, J.; McNamara, T.; Poikola, A.; Pollock, P.; Tait, J.; and Zijlstra, T. 2009. Open data handbook.

Donahue, J.; Jia, Y.; Vinyals, O.; Hoffman, J.; Zhang, N.; Tzeng, E.; and Darrell, T. 2014. Decaf: A deep convolutional activation feature for generic visual recognition. In *Proceedings of the international conference on machine learning*, 647–655.

Fukushima, K., and Miyake, S. 1982. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern recognition* 15(6):455–469.

Han, Q.; Zhu, Y.; Zeng, L.; Ye, L.; He, X.; Liu, X.; Wu, H.; and Zhu, Q. 2015. A road hotspots identification method based on natural nearest neighbor clustering. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 553–557. IEEE.

Jenks, G. F. 1967. The data model concept in statistical mapping. *International yearbook of cartography* 7(1):186–190.

Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; and Darrell, T. 2014. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, 675–678. ACM.

Karayev, S.; Trentacoste, M.; Han, H.; Agarwala, A.; Darrell, T.; Hertzmann, A.; and Winnemoeller, H. 2013. Recognizing image style. *arXiv preprint arXiv:1311.3715*.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *Nature* 521(7553):436–444.

LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; and Jackel, L. D. 1989. Backpropagation applied to handwritten zip code recognition. *Neural computation* 1(4):541–551.

LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–2324.

Leidig, M.; Teeuw, R. M.; and Gibson, A. D. 2016. Data poverty: A global evaluation for 2009 to 2013-implications for sustainable development and disaster risk reduction. *International Journal of Applied Earth Observation and Geoinformation* 50:1–9.

MacQueen, J., et al. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, 281–297. Oakland, CA, USA.

Maxfield, M. G. 1999. The national incident-based reporting system: Research and policy applications. *Journal of Quantitative Criminology* 15(2):119–149.

Miaou, S.-P.; Song, J. J.; and Mallick, B. K. 2003. Roadway traffic crash mapping: a space-time modeling approach. *Journal of Transportation and Statistics* 6:33–58.

Oquab, M.; Bottou, L.; Laptev, I.; and Sivic, J. 2014. Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1717–1724.

Razavian, A.; Azizpour, H.; Sullivan, J.; and Carlsson, S. 2014. CNN features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 806–813.

Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; and Rabinovich, A. 2015. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

United Nations. 2015. *Transforming our world: the 2030 Agenda for Sustainable Development*. United Nations.

World Health Organization. 2015. *Global status report on road safety 2015*. World Health Organization.

Xie, Z., and Yan, J. 2013. Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach. *Journal of transport geography* 31:64–71.

Zhou, B.; Lapedriza, A.; Xiao, J.; Torralba, A.; and Oliva, A. 2014. Learning deep features for scene recognition using places database. In *Advances in neural information processing systems*, 487–495.