SDCAE: Stack Denoising Convolutional Autoencoder Model for Accident Risk Prediction via Traffic Big Data

Chao Chen[1,2], Xiaoliang Fan[1,2,*], Chuanpan Zheng[1,2], Lujing Xiao[1,2], Ming Cheng[1,2], Cheng Wang[1,2]
[1] Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University, Xiamen, China
[2] Digital Fujian Institute of Urban Traffic Big Data Research, Xiamen University, Xiamen, China
Email: chenchao@stu.xmu.edu.cn, fanxiaoliang@xmu.edu.cn, {zhengchuanpan, 22920142203933}@stu.xmu.edu.cn, {chm99, cwang}@xmu.edu.cn

***Abstract*—Traffic accident is considered as one of main causes for traffic congestion in cities. There are many causal factors that may give rise to traffic accidents, e.g. driver characteristics, road conditions, traffic flows and weather conditions, etc. Due to uncertain factors as well as the contingency of accident occurrences, it is very difficult to predict traffic accidents. Many existing works have utilized classical prediction models to predict the risk of accidents on highways or road segments. However, predicting the risk of citywide accidents remains an open issue. To address this problem, we propose SDCAE, a novel Stack Denoise Convolutional Auto-Encoder algorithm to predict the risk of traffic accident in the city-level. First, we divided the city into regions by counting the number of accidents and traffic flows in each region. Second, we employed a deep model of stack denoise convolutional autoencoder which considers spatial dependencies to learn the hidden factors in accidents. Third, we conducted extensive experiments on two real-world cross-domain traffic big datasets from a major city of China for accident risk prediction. Experimental results demonstrate that SDCAE could outperforms five baseline methods.**

***Keywords*—traffic accident; convolutional autoencoder; risk prediction**

I. INTRODUCTION

With the popularity of vehicles and the increase of the number of drivers, there are more and more cars on the road. Once a traffic accident occurs, it is likely to cause road congestion. Nowadays, traffic accident has become one of the most serious urban challenges in big cities. It could not only cause economic loss, but also give rise to traffic congestion. For decades, traffic accident become an urgent problem to be solved for traffic polices. Many recent works have been done to reduce traffic accident, e.g., prohibiting lane change, setting the traffic light at intersection during rush hours, etc. In addition, to handle with traffic accident promptly and decrease the impact of traffic accident on road, they also try to monitor real-time traffic accident in the city level by calculating the change of traffic flows. However, due to the lack of specific devices, complex road networks and limitation of the monitoring algorithm, the result of monitoring is very poor, which might regard a traffic congestion for an accident. In general, the risk prediction of traffic accidents remains a challenging issue.

It is widely known that the occurrence of accidents is of contingency, which might make it extremely difficult in predicting accident occurrences. Fig. 1 shows that accidents occur in periodical patterns in statistics while happen unexpectedly case-by-case. For one thing, there are two regular peaks for accident occurrences (left in Fig. 1), which is corresponding to morning and evening rush hours. For another thing, there is a huge variance in accident occurrences if we compare of morning peaks of two continuous weekdays (top right vs. down right in Fig. 1). Early works [4-6] tend to focus on accident prediction on highway or road segments, with statistics and machine learning methods [15, 16, 18, 19]. Specifically, many works have been done on the risk prediction in traffic accidents, such as exploring causal factors (driver characters, road conditions, traffic flows and weather conditions, etc.) in traffic accidents [1-3]. However, correlations among those factors are very complicated. More recently, deep learning methods [21] have been employed to optimize the prediction model, by making use of the strong learning ability of deep neural networks to capture the correlations among causal factors of accident occurrences.

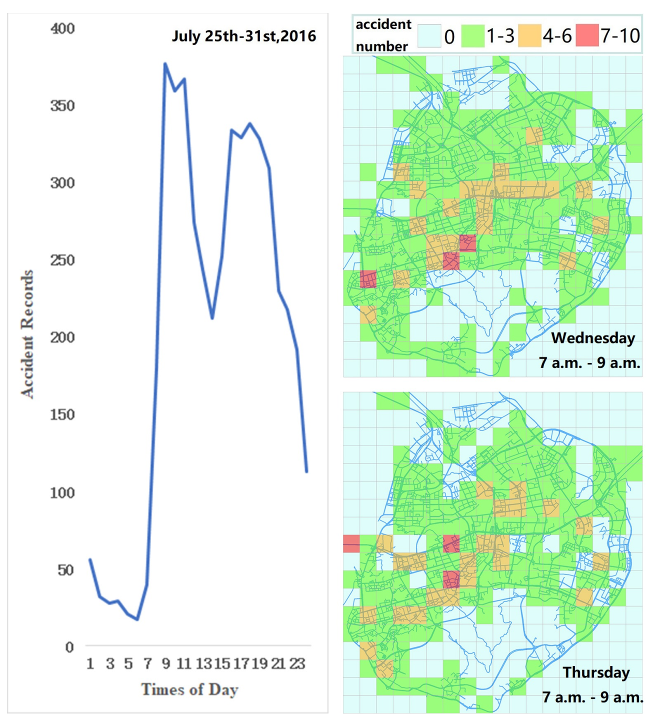

Figure 1. Spatio-temporal distribution of traffic accident

Inspired by studies on the accident risk prediction with deep learning methods [7, 20], we aim to make the best use of the strong learning ability of deep learning algorithms to uncover

978-1-5386-8034-6/18/$31.00 ©2018 IEEE
DOI 10.1109/CBD.2018.00065

the hidden dependencies of causal factors in citywide accidents prediction. However, it is a non-trivial task to predict the risk of citywide accidents, and we need to address the following two challenges:

- How to extract hidden features from accident data as well as cross-domain traffic data. The causal factors of accident occurrences are very complex that there are many external factors (i.e., dynamic traffic flows) could significantly affect the prediction model.
- How to integrate the influence of spatial dependences. Citywide traffic is dynamic with large-scare and network-wide features. Thus, it is urgent to consider the spatial correlations when predicting the risk of accidents.

To address challenges above, we propose *SDCAE*, a novel Stack Denoise Convolutional Autoencoder algorithm to predict the risk of traffic accident in the city-level. First, we preprocessed multi-sourced datasets to remove redundant and incorrect records. In addition, we divided the city into regions and count the number of traffic accidents and traffic flows in each region. Second, we employed a deep model of stack denoise convolutional autoencoder which considers spatial dependencies to learn the hidden factors in accidents through the addition of a convolutional layer. Third, we conducted extensive experiments on two real-world cross-domain traffic big datasets from a major city of China for accident risk prediction. Two datasets are: 1) 76,000 records of accident occurred in Xiamen from January 1st to August 31st, 2016; and 2) 1.7 billion of vehicle license plate recognition (VLPR) records in the city. Experimental results demonstrate that SDCAE could outperforms five baseline methods.

The remaining of this paper is organized as follows. Section II summarizes related works. Section III introduce our data. Section IV discusses the preliminary and propose our model. Section V describes the experiments and evaluations. Finally, Section VI concludes this work.

II. Related Works

Numerous of research about traffic accidents predictions has been done before. In early work, many researchers have done lots of research on finding what cause traffic accidents. A lot of studies showed that increased *speed* is associated with more accidents or higher accident rates [8-10]. Haynes *et al.* [11] concluded that *road curvature* has an inverse relationship with fatal crashes in urban settings. Dickerson *et al.* [12] examined the relationship between road traffic accident and *traffic flow* in London and found that a strong negative accident externality was associated with high traffic flows. Bedard *et al.* [13] discussed the impact of human factors, such as gender, age, driving years, on fatal injuries during the crash. Bergel-Hayat et al. [14] aims to highlight the link between weather conditions and road accident risk at an aggregate level. Li *et al.* [1] investigated the multiple factors associated with geometric factors (i.e., curve safety), such as weather (i.e., fog) and driver behavior (i.e., non-professional drivers, gender). Xi *et al.* [2] proposed AHP-Apriori algorithm and get a conclusion that the main influence factors of traffic accident are: driving experience, overload or not, road condition, weather conditions, etc. In addition, there is strong correlation between environmental factors and the accident types.

Furthermore, there are also many researchers trying to predict the risk related to traffic accident. Traditional crash prediction models employed *univariate models*, such as Poisson regression model, to explore the effects of roadway geometric factors on crash counts. Later, to address the over-dispersion issue of Poisson model, the Univariate Poisson Lognormal (UPLN) [15] and Negative Binomial (NB) [16] regression models were introduced to predict the total crash counts or crash counts by crash type. However, the major limitation of models mentioned above is that they ignore the correlations among different crash types or severities, which might result in biased parameter estimation and reduce model accuracy. Yuan *et al.* [17] investigated the problem of traffic accident prediction using heterogeneous urban data. Nowadays, with the development deep neural networks, these methods are used for traffic accident prediction. Ogwueleka et al. [18] proposed an artificial neural network model for road accident prediction. Jadann et al. [19] also using artificial neural network approach through analyzing the relationship between accidents and parameters affecting them for which data were available. Lu *et al.* [20] propose a model based on convolutional neural network to predict the traffic accident in highways. But experiment of the last three models are relative simple. Chen et al. [7] developed a deep learning model by autoencoder to make a risk prediction for traffic accident by using human mobility which is a first attempt to estimate traffic accident risk in a city level. In summary, our work differs from the-state-of-the-art methods that we extract hidden features from both accident data and two cross-domain traffic datasets, using a deep model of stack denoise convolutional autoencoder.

III. Data Description

A. Traffic accident data

There are about 350 records in Xiamen island every day, and around 10,000 records each month. After removing those redundant and incorrect data, we have collected about 76,000 records of traffic accident occurred in Xiamen from January 1st to August 31st, 2016. Each record includes accident ID, occurrence location, occurrence time, longitude, latitude and accident type. Table I introduces the sample of accident records.

Table Ⅰ. Sample of Traffic Accident records

ID	Location	Time	Longitude	Latitude	Accident Type
121432	SiBei road	2016/3/1 0:07:00	118.111232	24.47165	Crash
121565	TaiWan road	2017/3/1 8:03:00	118.020392	24.50132	Scratch
215154	HuYuan road	2016/3/1 9:17:00	118.107666	24.52879	Rear-end

B. *Traffic flow data*

We collect the traffic flow data by vehicle license plate recognition (VLPR) sensors. Recently, VLPR sensors are widely used for traffic monitoring. There are more than 240 VLPR sensors deployed in Xiamen island. The VLPR sensor will record every vehicle's information every three seconds. There are more than 7 million records generate every day. We have collected more than 1.7 billion of vehicle records from January 1st to August 31st, 2016, each record has attributes including Device ID, Driving Direction, Plate Number, Lane Number, and timestamp. Furthermore, there also has the VLPR sensors information including Device ID, Device Type, Device Location, Longitude, and Latitude. Table II introduces the sample of VLPR records for passing vehicles.

TABLE Ⅱ. SAMPLE OF VLPR RECORDS FOR PASSING VEHICLES

Device ID	Direction	License plate No.	Color	Lane No.	Time Stamp
311068	2A	A00001	2	2	1459440003
310043	2B	A00002	1	1	1459440009
311201	2A	A00003	2	2	1459440013

For privacy issues, the traffic related dataset doesn't include any confidential information. The data of vehicle number were hashed. In addition, the traffic related dataset we used in this work was authorized by Xiamen traffic police department.

IV. THE SDCAE MODEL

Recently, most research that inferring traffic accident are predicting whether it will happen or not. However, it is hard to predict whether there will have traffic accident or not. Since the factors that affect traffic accident are complex, most of them could not be observed in advance and the driver features are more influenced by subjective factors. Thus, in our research, we decide to infer that the risk of traffic accident instead of whether traffic accident will happen or not.

In our research, we merge traffic flow data, traffic accident data and time for the input of our model.

A. *Preliminary*

Before we build our model, we first change the data structure to appropriate our model that a matrix of data structure is needed.

a) Grid division: We combined the road network of Xiamen island and the distribution of VLPR devices. We divide the Xiamen island into an I×J region.

b) Traffic flows preprocess: We first map the VLPR devices into grid area by device coordianates. Then, we divide one day into 24-slices, and count the car records of every VLPR device every hour. If there are not any VLPR device in the grid, we assume that the traffic flow is zero. If there are only one device in the grid, the traffic flow is what the VLPR device in the grid count. If there are multiple devices in the same grid, we remove the duplicate records in a short-time intervals. We use $F_{(i,j,t)}$ to stand for the traffic flow at t moment in region (i,j). Once we count every hour every region traffic flow, we will normalized the data.

c) Traffic accident preprocess: We first extract useful attributes of data and remove the duplicate and redundant data. Then, mapping every accident record to the grid by coordinate and counting that how many accidents that every hour every grid haved. We use $S_{(i,j,t)}$ to stand for the number of traffic accident at t moment in region (i, j).

d) Time preprocess: We make another attribute by preprocess time. If it is daytime, make it 1, else make it 0. Since the light can also affect the risk of traffic accient.

B. *Proposed Framework*

In this module, we will show our framework in figure 2. As you can see, we collect multi-source data, and then preprocess the data to make them Mesher. Next, we will train deep learning model for feature extracting. After, the risk prediction model completed and we can use the model to predict traffic accident risk by real-time traffic data.

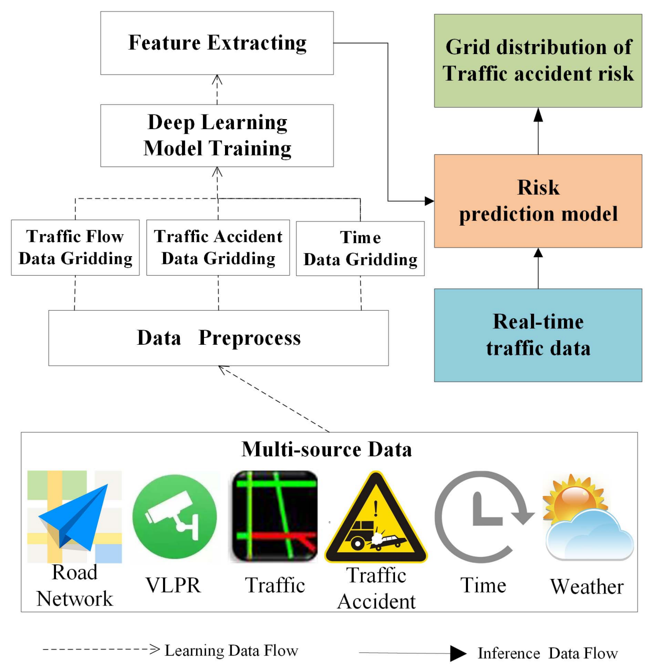

Figure 2. The overall framework

C. *Risk Prediction Model*

Vehicles could move one region to another in a short time, and in one hour they can move a long distance. As we know, the more adjacent the region, the stronger the association of traffic flow. Since the factors that affect traffic accidents are complex and traffic flows will affect near one region not just single region. Therefore, we plan to employ Stack Denoise Convolutional Autoencoder (SDCAE), which is a deep learning architecture that can extract hierarchical feature of input data. Our model combines the advantages of convolutional network and stack noise autoencoder. We know that the traffic flow is more closely localized in space and convolutional network can show this kind of characteristics. We build our model by using autoencoder (CAE). Then adding some noise into the input data to make the model more robust. Finally, we stack the convolutional autoencoder to make it have better performance than convolutional autoencoder.

When we train the model, firstly, we fuse the traffic accident data, traffic flow data and time data as input data $X_{(i,j,t)}$. Secondly, we add noise to the input data as new input $X'_{(i,j,t)}$, and we use $X'_{(i,j,t)}$ as input and $X_{(i,j,t)}$ as target. Thirdly, we train the convolutional autoencoder, the encoder of the model is as follow:

$$H_{1\_1}^{k_{1\_1}} = \sigma_{1\_1}\left(W_{1\_1}^{k_{1\_1}} X' + B_{1\_1}^{k_{1\_1}}\right) \quad (1)$$

$$H_{1\_2}^{k_{1\_2}} = \sigma_{1\_2}\left(W_{1\_2}^{k_{1\_2}} H_{1\_1}^{k_{1\_1}} + B_{1\_2}^{k_{1\_2}}\right) \quad (2)$$

where σ is the active function and W is a weight matrix present convolutional kernel and B is bias vector. Fourthly, we using the deconvolution reconstruct the $H_{1\_2}^{k_{1\_2}}$ as the decoder of the model, here is the equation:

$$H_{1\_3}^{k_{1\_3}} = \widehat{W}_{1\_3}^{k_{1\_3}} H_{1\_2}^{k_{1\_2}} + B_{1\_3}^{k_{1\_3}} \quad (3)$$

$$H_{1\_4}^{k_{1\_4}} = \widehat{W}_{1\_4}^{k_{1\_4}} H_{1\_3}^{k_{1\_3}} + B_{1\_4}^{k_{1\_4}} \quad (4)$$

where $\widehat{W}$ is the kernel of deconvolution. Finally, we convolutional the decoder layer of $H_{1\_4}^{k_{1\_4}}$ and assigned the value to Y.

$$Y_1 = \sigma_1\left(W_1^{k_1} H_{1\_4}^{k_{1\_4}} + B_1^{k_1}\right) \quad (5)$$

After that, we use the MSE as the loss function and calculate the MSE between Y and $X_{(i,j,t)}$. The equation of MSE is as following:

$$MSE = \frac{1}{s} \times \frac{1}{I \times J} \sum_{i=1}^{s} \left(X_i - \widehat{X}_i\right)^2 \quad (6)$$

where s stand for the sample of input, I and J stands for the rows and columns. Until now, the denoise convolutional autoencoder (DCAE) has been trained. DCAE is based on autoencoder. The different between the autoencoder and DCAE are that the noise was added into train samples which make the model more robust, the fully connected layers were changed into the layer of convolution or deconvolution. As we know, traffic flows have more influence in neighboring region. Therefore, DCAE is more suitable than autoencoder in our research.

When we try to build SDCAE model, we need to train every single DCAE. Once we have trained every DCAE, we extract the encoding layers of each DCAE and stack them to make a deep network by feeding the extracting feature of every DCAE to build the model of SDCAE. We use SDCAE for reconstructing the input. After that, we go onto the stage of risk prediction. SDCAE is an unsupervised network, to use SDCAE for prediction, we need to add a predictor on the top of the network to make the network into a supervised network.

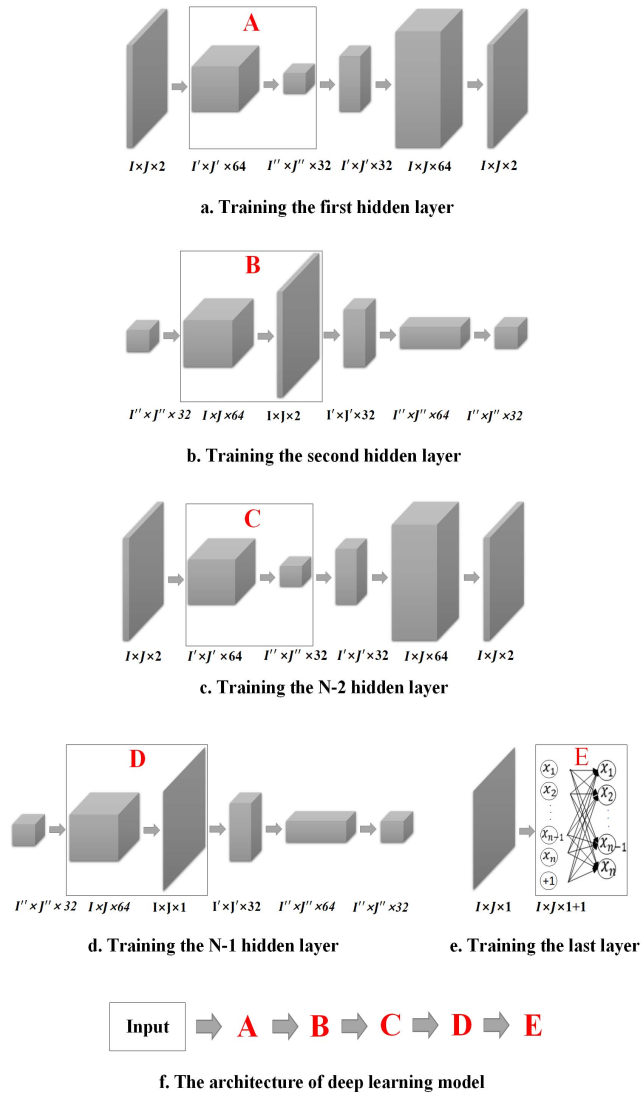

Figure 3. The Training of Our Model. First, we training each individual model from Fig. 3a to Fig. 3e, the output of each layer is used as the input of the next layer. Then we select the encoder layer that in the frame and stack them together to be our model as shown in Fig. 3f.

As follows, we fuse the multi-source data as input $X_{(i,j,t)}$, and use the accident data $S_{(i,j,t)}$ as label. The procedure is as follows:

a) Train the first DCAE which consist of two convolutional layers, two deconvolutional layers and one convolutional layer by minimizing reconstruction error. The first two layers of DCAE are the encoder part and we extract encoder layers.

b) Taking the former output as input and train the second DCAE which consist of two deconvolutional layers, two convolutional layers and one convolutional layer. In the second DCAE, it is a little different from the first DCAE where swap the training order of convolutional layers and deconvolutional layers.

c) Iterate step a) and b) in turn for desired number of layers.

d) Using the output of step c) as input into a supervised layer for risk prediction.

e) Fine-tuning the parameters using training data and make the model more suitable for our data.

V. Experiments and Evaluations

A. Experimental Settings

We collect more than 1.7 billion passing records from more than 240 VLPR devices and more than 76000 records of accident from January 1st to August 31st, 2016. To evaluate the performance of our model, we split our data to training set and test set, the training set is 80% of all data, the remaining for testing and evaluation. In our experiment, we fuse multi-source data as different channel of convolutional layer. We totally train four single DCAE and make them to be SDCAE network, then add a logistic regression for risk prediction. The experiment developed with python3.5 and tensorflow1.7, and we use GPU node on Argon with NVIDIA Tesla P100 Accelerator Cards.

B. Evalution Metrics

In order to evaluate the accuracy of our model, we evaluate the prediction error of our model with two different metrics, including mean absolute error (MAE) and mean relative error (MRE). They are separately defined as follows.

$$MAE = \frac{1}{s} \times \frac{1}{I \times J} \sum_{n=1}^{s} \sum_{i=1}^{I} \sum_{j=1}^{J} |X_{nij} - \widehat{X_i}| \quad (7)$$

$$MRE = \frac{1}{s} \times \frac{1}{I \times J} \sum_{n=1}^{s} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{|X_{nij} - \widehat{X_i}|}{X_{nij}} \quad (8)$$

C. Baselines

We have selected several models for the baseline models.

- *Logistic regression (LR)*: Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function.
- *Random Forest (RF)*: Random Forest is an ensemble leaning method for classification, regression and other tasks.
- Decision Tree (DT): Decision Tree is to create a model that predicts the value of target variable based on several input variables.
- *Linear Regression (LN)*: Linear regression is a linear approach to modeling the relationship between a dependent variable and independent variables.
- *Stack denoise autoencoder (SDAE)*: A SDAE model is stack by autoencoder, it is an unsupervised learning and can well reconstruct the input, then use the reconstruct data to make prediction.

D. Performance Evaluation

To evaluate the performance of our model, we have compared the metrics with the baseline model. In the model, we train them and select the best performance of our training in these models. As a result, the smaller the error is, the better the results are. We show the results as follows.

Comparison with Baselines. As shown in Table III, the deep learning methods outperforms traditional machine learning algorithms in general. Since the correlations among traffic accidents causal factors are complex, the traditional machine learning algorithms have limitation in uncovering the hidden correlations of causal factors.

TABLE III. PERFORMANCE EVALUATION

Algorithm	MAE	MRE
LR	1.120	0.996
RF	1.095	0.970
DT	1.081	0.939
LN	0.192	0.929
SDAE	0.115	0.879
SDCAE (4 hidden layers)	0.098	0.865
SDCAE (8 hidden layers)	0.095	0.848
SDCAE+BN (8 hidden layers)	**0.092**	**0.796**

Comparison with SDAE. The SDCAE is different from SDAE where the connections among layers are convolutional rather than fully-connected. In general, the prediction effect of SDCAE is obviously better than that of SDAE. The main reason is that convolutional autoencoder can better handle spatial dependencies on traffic flows than autoencoder itself.

Variants of SDCAE. SDCAE+BN (batch normalization) with 8 hidden layers perform the best among all three SDCAE variants. Specifically, BN is a method which could normalize the input to a distribution with a mean value of 0 and a variance of 1. Such normalization might increase the accuracy of the model and can prevent overfitting. In addition, BN could ensure that the distribution of input data on each layer is stable which could speed up the training process. Furthermore, the reason why 8 hidden layers perform better than that of 4 hidden layers is that the depth of hidden layers is corresponding to the causal factors to be learned.

VI. Conclusion

We proposed a model to predict the risk of city level from a new perspective. In our experiment, we combined with the distribution of road network and VLPR devices in Xiamen island divide Xiamen into grid $I \times J$. We conducted extensive experiments on two real-world cross-domain traffic big datasets from a major city of China for accident risk prediction. Comparing our model with baseline models, we found that our model performs better than the other baseline models.

In the future, our work can be extended. For instance, we will try to fuse other external datasets that may have impact on the risk of traffic accidents (i.e., weather conditions) to further train our model.

ACKNOWLEDGEMENT

The work is supported by grants from the Natural Science Foundation of China (61300232); and the China Postdoc Foundation (2015M580564). The corresponding author is Xiaoliang Fan.

REFERENCES

[1] X. Li, X. Yan, S. C. Wong, "Effects of fog, driver experience and gender on driving behavior on S-curved road segments, " Accident Analysis and Prevention, vol. 77, Apr. 2015, pp. 91-104, doi: 10.1016/j.aap.2015.01.022.

[2] J. Xi, Z. Zhao, W. Li, Q. Wang, "A traffic accident causal analysis method based on AHP-apriori," Procedia engineering, vol. 137, July. 2016, pp. 680-687, doi: 10.1016/j.proeng.2016.01.305.

[3] R. Yu, M. Abdel-Aty, M. Ahmed, "Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors," Accident Analysis & Prevention, vol.50, Jan. 2013, pp. 371-376, doi: 10.1016/j.aap.2012.05.011.

[4] S. Park, S. Kim, Y. Ha. "Highway traffic accident prediction using VDS big data analysis," The Journal of Supercomputing, vol. 72, Jul. 2016, pp. 2815-2831, doi:10.1007/s11227-016-1624-z.

[5] M. Yasin Çodur, A. Tortum, "An Artificial Neural Network Model for Highway Accident Prediction: A Case Study of Erzurum, Turkey," PROMET-Traffic & Transportation, vol. 27, 2015, pp. 217-225, doi: 10.7307/pttv27i3.1551.

[6] X. Meng, L. Zheng, Y. Shi, "Research on accident prediction models for freeways considering effect of lane change maneuvers," Journal of highway and transportation research and development, vol. 31, 2014, pp. 121-126, doi:10.1109/ICMTMA.2015.205.

[7] Q. Chen, X. Song, H. Yamada, R. Shibasaki, "Learning Deep Representation from Big and Heterogeneous Data for Traffic Accident Inference," AAAI. 2016, pp. 338-344.

[8] C. Wang, M.A. Quddus, S.G. Ison, "The effect of traffic and road characteristics on road safety: A review and future research direction," Safety science, vol. 57, Aug. 2013, pp. 264-275, doi: 10.1016/j.ssci.2013.02.012.

[9] G. Nilsson, "Traffic safety dimensions and the power model to describe the effect of speed on safety," Bulletin-Lunds Tekniska Högskola, Inst för Teknik och Samhälle, Lunds Universitet, 2004, pp. 221.

[10] L. Aarts, I. Van Schagen, "Driving speed and the risk of road crashes: A review," Accident Analysis & Prevention, vol. 38, Mar. 2006, pp. 215-224, doi:10.1016/j.aap.2005.07.004.

[11] R. Haynes, A. Jones, V. Kennedy, I. Harvey, T. Jewell, "District variations in road curvature in England and Wales and their association with road-traffic crashes," Environment and Planning A, vol. 39, May 2007, pp. 1222-1237, doi:10.1068/a38106.

[12] A. Dickerson, J. Peirson, R. Vickerman, "Road accidents and traffic flows: An econometric investigation," Economica, vol.67 Feb. 2000, pp. 101-121, doi: 10.1111/1468-0335.00198.

[13] M. Bedard, G. H. Guyatt, M. J. Stones, J.P. Hirdes, "The independent contribution of driver, crash, and vehicle characteristics to driver fatalities," Accident Analysis & Prevention, vol. 34, Nov. 2002, pp. 717-727, doi: 10.1016/S0001-4575(01)00072-0.

[14] R. Bergel-Hayat, M. Debbarh, C. Antoniou, G. Yannis, "Explaining the road accident risk: weather effects," Accident analysis and prevention, vol. 60, Nov. 2013, pp. 456, doi:10.1016/j.aap.2013.03.006.

[15] D. Lord, F. Mannering. "The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives," Transportation Research Part A: Policy and Practice, vol. 44, Jun. 2010, pp. 291-305. doi: 10.1016/j.tra.2010.02.001.

[16] V. Shankar, F. Mannering, W. Barfield. "Effect of roadway geometrics and environmental factors on rural freeway accident frequencies," Accident Analysis & Prevention, vol. 27, Jun. 1995, pp. 371-389. doi: 10.1016/0001-4575(94)00078-Z.

[17] Z. Yuan, X. Zhou, T. Yang, J. Tamerius, R. Mantilla, "Predicting traffic accidents through heterogeneous urban data: A case study", 6th International Workshop on Urban Computing (UrbComp 2017), Aug. 2017.

[18] F.N. Ogwueleka, S. Misra, T.C. Ogwueleka, L. Fernandez-Sanz, "An artificial neural network model for road accident prediction: a case study of a developing country," Acta Polytechnica Hungarica, vol. 11, 2014, pp. 177-197.

[19] K.S. Jadaan, M. Al-Fayyad, H.F. Gammoh, "Prediction of road traffic accidents in jordan using artificial neural network (ANN)," Journal of Traffic and Logistics Engineering, vol. 2, June 2014, doi: 10.12720/jtle.2.2.92-94.

[20] L. Wenqi, L. Dongyu, Y. Menghua. "A model of traffic accident prediction based on convolutional neural network," Intelligent Transportation Engineering (ICITE), 2017, pp. 198-202, doi: 10.1109/ICITE.2017.8056908.

[21] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT Press, 2016.