Stack ResNet For Short-term Accident Risk Prediction Leveraging Cross-domain Data

Zhengyang Zhou*, Lianliang Chen, Chaochao Zhu, Pengkun Wang
University of Science and Technology of China, Hefei, China
Email:{zzy0929*, cll006, cczhu, pengkun}@mail.ustc.edu.cn

***Abstract*—With an increasing number of private cars, traffic accident prediction plays an important role in urban management. Existing methods tend to utilize classic learning techniques with historical observations on accident records. Therefore, they fail to consider cross-domain data into consideration and ignore the spatiotemporal dependencies. Recently, as more heterogeneous urban data are available, it is promising to predict citywide accidents in a more fine-grained way. However, it is challenging due to the inherent sparsity of traffic accident records and the complexity of multiple influenced factors. Besides, real-time data cannot be collected completely for short-term forecasting. In our paper, we propose a multi-view spatiotemporal deep learning framework to fuse cross-domain urban data. Specifically, we first assign heterogeneous data into the corresponding grid, then extract the features related to accidents into tensors or vectors. Moreover, we propose a ResNet based multi-task learning framework with speed inference model to realize the prediction of near future's accident risk distribution.**

***Index Terms*—traffic accident prediction, ResNet, urban computing, spatiotemporal data mining**

I. Introduction

Traffic accidents usually result in large cost of property and even people's life. Thus, there are increasing requirements for citywide traffic accidents forecasting, which provides advice for the traffic police force arrangement in different blocks previously and safe route recommendation. Plenty efforts have been made on this issue based on historical accidents record, only few consider weather records or cross-domain data [1], [3], [4]. Although some deep learning methods enable promising results, their works have several shortcomings more or less. They applied data only from one point of view and failed to take the periodical patterns into consideration and overlooked the dependency of cross-domain urban data, some just predicted the limited scale which is lack of generalization, or proposed a data-hunger model which needs a large scale of data to feed in [3]. With the development of new techniques of collecting cross-domain urban data, this task remains challenging because of the sophisticated relationship between them. When compared to the traffic flow prediction task, accidents risk prediction is more difficult [4]. This is mainly because accidents in the road network are usually caused by complex factors and large variant dependencies are not available every time, such as the real-time speed. Additionally, some random human factors also contribute to accidents. Nevertheless, it does not mean the accident risk distribution is unpredictable. We argue that it can be feasible for the following reasons. First of all, more urban data collected by sensors is available for researchers and new techniques of deep learning are developed. Secondly, the risk of accidents at one certain place tends to follow the variation of traffic volume and be influenced by auxiliary factors. Thus, a sequence of past observations of events can imply many possible futures. Third, unreasonable road structures may lead to frequent traffic accidents. The goal of our task is to predict the near future accident risk by mining the correlation between historical accident records and cross-domain urban data. According to the previous research, the occurrence of accidents is correlated with weather, road network structure, density of population [3], [4], [7]. In our research, the data we collect includes accident records with spatial and temporal information, meteorological data, and other types of traffic-related data. All these types of data are preprocessed by grid-based method. They provide contextual information including both spatial and temporal dimensions, allows us to have a better understanding of dynamic traffic accidents distribution. To handle the complex dependency and limitations aforementioned, we propose a novel framework named Stack ResNet for hour-level Accident Prediction (SRAP) and model the dynamic spatiotemporal features from multiple temporal views, making the accidents prediction globally and simultaneously. To summarize, the main contributions of our work are four-fold:

- A variant of Convolution Neural Network (CNN) based multi-task deep learning framework is proposed to handle both static and dynamic urban data for accident risk distribution forecast.
- Spatiotemporal patterns of traffic accidents are mined and three ResNet modules share the same structure are expected to learn the representation of spatiotemporal features from three temporal views, which implies the closeness, trend and period of accidents variation.
- We construct a new citywide speed inference model which depends on spatial features of the road network. Further, by investigating the relationship between accident counts and dynamic traffic conditions, we propose the multi-task learning framework, leaving prediction of speed and strength of human mobility as two auxiliary tasks.
- Experimental results on real-world datasets demonstrate our proposed method outperforms other alternative models and the training time significantly reduced by 10 times with LSTM cell dropped when compared to ConvLSTM.

978-1-7281-4094-0/19/$31.00 ©2019 IEEE

To be specific, we also present a case study of accident prediction results in one day and give an intuitive comparison between the prediction results and ground truth.

The remainder of this paper is organized as follows. We first give some preliminaries including problem definition, ingredients of features and deep residual network in Section 2. Then we overview spatiotemporal patterns of traffic accident distribution and elaborate our proposed ST-MGAP framework in Section 3. Then extensive experiments are conducted and competitive results are demonstrated in Section 4. Lastly, we conclude our paper as well as discuss the weakness of our method and future work in Section 5.

II. Preliminaries

In this section, we give the general definition of our problem, and then introduce the data source we collect for our experiments. Last, the property and characteristics of ResNet block are presented.

A. Problem Definition

In this section, we briefly give the definitions and introduce the problem of accident prediction in the road network.

*Definition 1 (**Urban Region**): To perform fine-grained citywide accident prediction, we take grid-based method to divide the study area(NYC except for Staten Island) into disjointed square-shaped grids and each grid is not overlapped each other. As a consequence, the total number of regions is denoted as N.*

*Definition 2 (**Accidents Risk Records**): A two-dimension tensor denotes as R_{it}, where t is the time interval and i is the region index in the defined study area. To be noted that $R_{\cdot t}$ means citywide accident risk at time interval t.*

*Definition 3 (**Features**): Two traffic attributes, average speed and taxi trip records, are indicators of traffic conditions. They are jointly denoted as C_{it}. Similar to Definition 2, $C_{\cdot t}$ means citywide accident risk at time interval t. A set of static road network or traffic related features at time interval i are defined as $F_i = \{f_1, f_2, f_3, ..., f_n\}$.*

*Definition 4 (**Citywide Accident Risk Prediction**): Given a sequence of observed historical traffic accident records and traffic-related observations.*

$$\{R_{\cdot t}, C_{\cdot t}, F_t | t = 0, 1..., T\}$$

The problem is to predict citywide traffic accident risk distribution $R_{\cdot (T+1)}$.

B. Data Source

We collect accident-related cross-domain data between Jan 1, 2017 and Dec 31, 2017 in New York City (NYC Except Staten Island). More details are explained as follows:

- Road Network Data: Road Network structure plays a vital role in traffic condition prediction. Due to road lanes and road types have an interaction effect on road segment speed and road condition evaluation, we collect data of (1) Number of road lanes, and (2) Road types. With the truck proportion increasing, cars and coach tend to keep away from the truck, which leads to high variance of vehicle speed. Hence, we collect (3) Truck proportion. (4) Overhead electronic signs distribution is also be considered in our model because it is proven to be related to accidents especially at nights in practice [5].
- Meteorological Data: Accidents tend to happen when coming up with the extreme weather, such as road icy, heavy snow. Therefore, we integrate the following datasets: (1) Precipitation (2) Snowfall (3) Speeds of wind (4) Degree of temperature (5) Pressure for our prediction task.
- Social Data: During our experiment, we found that (1) demographic data has influences on traffic accident distribution. It is also reasonable that the higher density of people in a region, the higher probability to have accidents. (2) Point of Interest (POI) data which represents a special location tends to have a specific traffic pattern, with name, category and coordinates. (3) Human Mobility Data: We collect human mobility data such as taxi trips. Taxi trip records illustrate people's arrivals and departures in each grid, which help us capture the mobility pattern of urban citizens.
- Calendar Data: Accident counts vary periodically like crowd flows when during holidays or weekends (e.g. Chinese Spring Festival), it can be significantly different from the count during normal days. Thus we added the holiday and calendar data into our method.

C. Deep ResNet

Deep neural networks, especially CNN, have achieved a series of breakthroughs in challenging tasks such as image classification, object detection in Computer Vision (CV) fields. With the continuous development of urban computing, grid-based method is employed to citywide spatiotemporal data mining. Therefore, it is a significant milestone that uses CNN to traffic prediction task. Unfortunately, with the layers of network increasing, accuracy gets saturated, and then degrades quickly. This is mainly because of the gradient degradation, which is not caused by overfitting. Then Deep ResNet is proposed by He [6], and it outperforms other frameworks in CV datasets. The basic idea of the residual block is a residual mapping which creates a shortcut. It can be formulated as:

$$X_{l+1} = X_l + F(X_l) \tag{1}$$

This trick can combine the output of previous layers with the current layer and feed into the next layer to alleviate the gradient degradation.

In our research, in order to stack capture the distant potential spatial correlation and propagation patterns of traffic accidents, we introduce ResNet to our framework stack layers.

III. Proposed Framework: Stack ResNet for Traffic Accident Prediction

In this section, we demonstrate our SRAP framework for traffic accident prediction. We first analyze the spatial dependency for traffic patterns and formulate the study area as an image-like map. Then we elaborate on how to build our proposed framework with stack ResNet and large-scale cross-domain data, and an overview of the proposed SRAP is shown in Figure 1.

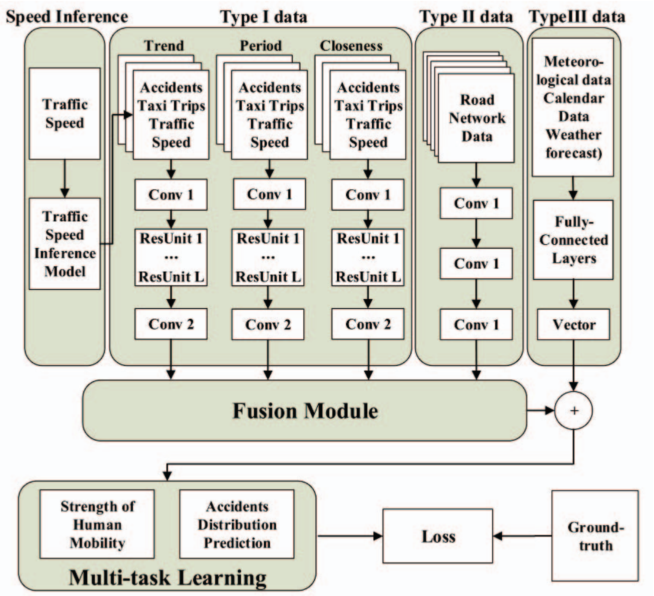

Fig. 1: Overview of proposed framework SRAP

A. An Analysis of Spatiotemporal Pattern

In this section, we analyze spatiotemporal patterns in traffic accidents. We first classify the datasets into several types: (1) Type I: Datasets including variables only varied within space: road network features. (2) Type II: Datasets including variables change with both space and time. (3) Type III: Datasets including variables only change with time. Specifically, the first types of datasets are number of road lanes, road sign distributions and demographic data. The second type of datasets are accident counts, average speed and taxi trips. The third type datasets are instantiated as weather and calendar data. The first two types of datasets are aggregated into the homologous grids and the third type datasets will be encoded as the fixed-length vectors.

On the one hand, the three temporal properties can be regarded as closeness, period, trend, because accident distribution follows the variation of tidal traffic flows. On the other hand, one region's accident counts can be influenced by its neighborhood propagation and even the traffic conditions of distant regions. For instance, accidents often induce traffic congestions and vehicle accumulations, which lead to frequent overtaking and relatively high variances of traffic speeds, then the possibility of accidents in the nearby regions becomes higher. Besides, the congestion will force drivers to alter the planned route then lead to the sudden increasing flows of other roads and propagation of the accidents risk. Furthermore, some types of traffic-related data are latent factors for crashes, for example, human mobility can be represented by the average speed and taxi trips and meteorological data will influence real-time urban conditions. Thus, all kinds of datasets are related and potentially contribute to the dynamic accidents risks in the road network.

B. SRAP Framework

Figure 1 shows that the framework of our spatiotemporal accident forecasting model, which consists of three components: (i) CNN based feature generator, (ii) Feature map fusion section, (iii) Accident forecasting based on feature sequence. Well-organized data flow of key timesteps will be fed into our framework, we will elaborate them in the following sections.

1) Feature Extraction: Firstly, we choose the study area and employ grid-based method to generate a grid-shaped map. The entire city is divided into 27×27 regions in total. Then all kinds of traffic-related data are assigned into the corresponding girds. For type I data, it is a spatial and temporal variant, we assign them into the correct grid, and then aggregate them into tensors with a stime axis. For type II, it is spatial variant but time-invariant, it will be stacked as a series image-like dataset. Last, Type III data will be aligned and converted into a fixed-size vector, and then stack in the time dimension. Overall, all kinds of data are transferred into tensors or encoded as numerical vectors.

2) Citywide Speed Inference Model based on Spatial Dependency: Motivated by the important role of vehicle speed in accidents prediction, it is necessary to model the citywide speed globally. We propose a spatiotemporal semi-supervised method by taking the advantages of both spatial dependencies. We assume that road traffic speeds can be relied on the following subregion-related casualties: (1) regions are close to each other can be with the similar traffic speed patterns and (2) those roads which are far from each other but have the same static road network structure and similar functionalities will possess the relevant average speed. Then we can establish our inference model with the weighted-regression formulation. We can generate vectors represent the corresponding region by concatenating the POI distributions with the road network features. The similarity of each region is defined as the variant of cosine distances of region description vectors D_i. The distance is defined in a Gaussian kernel. The affinity graph matrix M can be learned by the collected spatial features, formulated as:

$$D = [Road\,Features, \mathrm{POI}] \tag{2}$$

$$M = \begin{bmatrix} dist(D_0, D_0) & ... & dist(D_0, D_n) \\ ... & dist(D_i, D_j) & ... \\ dist(D_n, D_0) & ... & dist(D_n, D_n) \end{bmatrix} \tag{3}$$

$$dist(D_i, D_j) = \exp(-\frac{-\cos(D_i, D_j)}{2\sigma^2}) \tag{4}$$

where σ is the parameter to control the scale.

3) Proposed Model of SRAP: Based on feature tensors and vectors, we propose a Stack ResNet for Accident Prediction (SRAP) to predict the citywide accidents counts. According to classification of variables, we define the feature matrices as

A_1, A_2, A_3 to represent Accident records, Speed values and Taxi trip records respectively, B_1, B_2, B_3, B_4, B_5 represent Type II data (Road Network features), V_0 represents Type III data (meteorological and calendar data) for simplification. As a consequence, each group of Type I data becomes a series of image-like frames. Characteristics of temporal-dependent dynamics in road network are instantiated as trend, period and closeness. Inspired by it, features of type I data are extracted by three ResNet structures respectively. Therefore, training dataset A_1, A_2, A_3 are divided into groups of well-designed samples, each training sample has three groups of spatiotemporal features, representing the features of three properties, as well as one groud truth. Three modules which applied to extract features share the same network structure and model the three properties respectively from multiple temporal views. To be clearer, we take the first input as an example as Figure 2 shows. Each group stack l_w, l_p, l_c adjacent traffic records, where l_w, l_p, l_c denote the input length of frames in three properties. The adjacent records are deemed as multiple channels of one image, the intuition behind this idea is to capture the low-level feature correlations in time dimension. Then our task degenerates into the problem of predicting the next frame based on a series of key timesteps appearing before, it can be formulated as:

$$\begin{aligned}\{X_{t-24\times 7-l_w}, ..., X_{t-24\times 7}; X_{t-24-l_d},\\ ..., X_{t-24}; X_{t-l_c}, ..., X_{t-1} | X_t\}\end{aligned} \tag{5}$$

where X_t indicates the traffic attributes at the t_th interval. In addition, we utilize L Residual blocks to extract the spatiotemporal features, and eventually, a deconvolution layer is used to decrease the size of the feature map to fit the channel of output. The grow rate (number of filters) is set as 64 and 3 for Conv and DeConv layer. It is easier for the network to learn the residual error than original features due to the nice property of ResNet has. The static features of road network attributes can be assigned into the grids and extracted by a CNN network. The CNN structure not only takes the advantages of extracting features from neighborhood including congestion propagation, which probably leads to high-variance speed and increasing risks of intersections, also makes the best of fitting residual information. Type III data is encoded by two fully-connected layers and then is fused with spatiotemporal features. With the fusion mechanism, our model reweights the different temporal dependency, thus the model can dynamically aggregate the output of the three residual neural networks autonomously.

4) Multi-task Learning for Accident Distribution: The citywide accident records at one interval in a certain region are transferred into Risk strength by the formula:

$$Risk = \sum_i (Injure_i + Prov_i + 2 \times Fatal_i) \tag{6}$$

where $Injure_i$,$Fatal_i$ denote the number of persons injured and died at the i_th accident in the corresponding interval at a specific region. Then traffic speed, taxi trip records, which are indicators of strength of the citywide human mobility, are normalized by max-min normalization and scale to 0 to 1. As the human mobility has high correlations with accident risk, we adopt multi-task learning method for the training process in order to improve the extensiveness, then strength of human mobility and accidents risk will be trained and predicted synergistically. The intuition behind this mechanism is to enhance the representation and promote the prediction accuracy of acicdent risk. The loss function and the weight for auxiliary tasks will be introduced in the experiment setup.

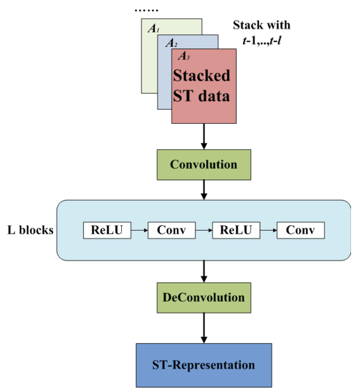

Fig. 2: One of stack ResNet blocks

$$Loss(\theta) = \lambda_1 \times mse\_acc + \lambda_2 \times mse\_Ttrip \tag{7}$$

To be noticed that, mse_acc, mse_speed, mse_Ttrip denote the mean square error(MSE) of the accident, speed and taxi trip.

IV. Experiments

In this section, we perform comprehensive experiments to evaluate our model on real-world NYC datasets.

A. Experiments Setup

1) Study Area: : The study area is NYC (except for Staten Island) and the number of the total regions is 729.

2) Datasets: : NYC datasets range from 1st. Jan to 31th. May, which can obtain by visiting NYC opendata [1] and Taxi & Limousine Commission [2]. All features we collect are assigned into the corresponding regions. As formulated in section 3.2.1, three types of data of Type I, II, III are preprocessed separately. For data split, we choose the last 20 days as the test sets. The total datasets are transferred into 3144 samples and each sample has 9 one-hour intervals, each interval includes 3 traffic attributes, accident risk, average speed as well as taxi trip records.

3) Platform and hyperparameters: : We perform our experiments on GPU Tesla V100-PCIE-16GB. Tensorflow and Keras libraries are involved to help build our deep neural network. To capture the propagation and effects of the surroundings, the kernel size of CNN is fixed and set as (3,3,3). Time interval of one day is fixed to 1 hour. The external factors

[1]https://data.cityofnewyork.us/
[2]https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

are converted into a fixed-length vector in the size of six. The length of closeness, daily period and weekly period is set to 3. The loss function is MSE. The residual blocks are set varying from 4 to 6. The weight between the main and auxiliary task is set as λ_1=1, λ_2=0.6 according to the grid search method. Due to the space limitation of this paper, we omit the details of the process for adjusting the parameters.

B. Verifications for Correlations between Accidents and Taxi trips

We present the hourly citywide traffic accident counts and taxi records in one figure, investigating the correlations between accidents and taxi trips. As can be observed in Figure 3, it reveals accident count reaches a peak at midnight and then a relatively higher risk peak appears at afternoon from the temporal dimension. This may imply that drivers tend to be tired at midnight and afternoon, as fatigue driving takes account for main reasons for crashes. Combined the taxi trip records together, we conclude that traffic accident risk is generally consistent with the taxi trip records, which can be indicators of the strength of citywide human mobility. This verifies that it is feasible to model the three properties together. There is also an interesting finding in our research, most citizens fall asleep at midnight and few people will attend parties or be engaged in transportation service, then they have a high possibility to take the risk due to the fatigue. In other words, either you don't go out, if you do, you will encounter greater danger at midnights.

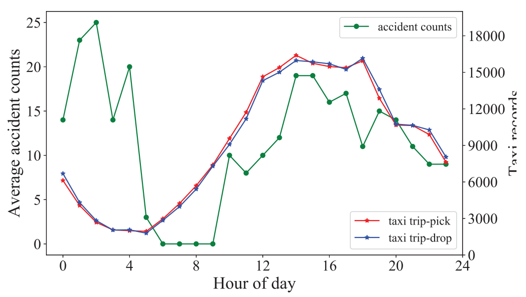

Fig. 3: Daily records of citywide accidents and taxi trips

C. Baselines

To evaluate the proposed framework comprehensively, we reproduce the algorithms put forward by Yuan [4], which is the state of the art method for accident prediction, and we also use a shallow model called Auto Regression Integrated Moving Average (ARIMA) method, which is considered as the classic model for sequence traffic prediction. Results can be found in Table 1. For fairness, the types of data are organized as our model except for ARIMA.

ARIMA: Auto-Regressive Integrated Moving Average (ARIMA), which is a classic machine learning algorithm, well-known for understanding and predicting future values, especially for time series predictions.

ConvLSTM-2 blocks: We set 16,1 filters of each ConvLSTM block. External factors are also fed into the model and fuse the ST features.

ConvLSTM-3 blocks: We use 16, 32, 1 filters of each ConvLSTM block.

D. Metrics and Results

1) Metrics: (i) Mean square error (MSE) and Root mean square error (RMSE) are the general metrics for regression tasks. Here we introduce the two metrics for our evaluation.
(ii) Accuracy: Since our model generates a risk map of the whole city, it is difficult for us to calculate the accuracy simply by employing the traditional accuracy definition. In this map, regions have higher risk tend to have a higher risk to be accidents in the near regions. To evaluate the efficiency of models, we compare the regions with high risk with the ground truth, to be specific, we select the regions which have highest risk of top-K as the area predicted to be an accident, where K is the adaptive and learnable threshold for different traffic conditions. The metric *Accuracy* can be formulated as:

$$\begin{aligned} Accuracy = \frac{1}{N}(\sum_i I(accPT_i = accTrue_i) + \\ \sum_j I(accPF_j = accNeg_j)) = \frac{TP+TN}{N} \end{aligned} \tag{8}$$

where $I(\cdot)$ is an indicator function, acPT, accPF denote the prediction results is high risk and low risk, and accTrue, accNeg denote the ground truth is occurrence and non-occurrence respectively.

(iii) Acc_detection_rate: The ratio of the regions predicted to be high-risk and the actual accident to the total number of real accidents. This means it is the real predicted results without any modification.

$$Acc\_detection\_rate = \frac{TP}{TP+FN} \tag{9}$$

2) Evaluation and Competitive Results: The performance of different models is presented as follows in Table 2. It reveals in Table I that deep models can easily beat the shallow machine learning model which does not involve the spatiotemporal context. The proposed method improves the Acc_detection_rate by nearly 20% compared to the best baseline (ConvLSTM-2blocks) from 19.62% to 23.38% and we achieve the highest accuracy. It is worth mentioning that our model not only outperforms alternative algorithms in accuracy, moreover, in terms of complexity of time, SRAP model is also significantly reduced. We drop the module of LSTM, which usually has a large cost of computation, and the ConvLSTM is trained for more than 1000 seconds while ours only takes 106 seconds. This implies we employ the fusion method without LSTM still can reach a quite equivalent performance and even better when data and model are well-designed.

E. Case study

Furthermore, we visualize prediction results in one selected day, as Figure 4 shows. The red circle in the picture reveals

TABLE I: Performance of different models

Model	MSE	RMSE	Accuracy	Acc_detection_rate	Training time (s)
ARIMA	1.6801	1.2759	84.24%	14.23%	96
ConvLSTM-2blocks	1.0300	1.0140	86.48%	19.62%	984
ConvLSTM-3blocks	0.8620	0.9284	87.47%	13.39%	>1000
ResNet-4blocks	**0.1600**	**0.4000**	**88.89%**	**23.38%**	**116**
ResNet-5blocks	0.1671	0.4086	88.20%	21.92%	127
ResNet-6blocks	0.2027	0.4494	88.30%	20.27%	135
ResNet-4blocks(single-task)	0.1680	0.4098	86.27%	21.88%	102

selected areas for comparison. Different depths of color represents different levels of traffic risk in different regions and darker color means higher accident risk.

As can be observed in Figure 4, we select three typical intervals that have a high frequency for accident occurrence in one day, at 1 a.m., 8 a.m., 9 p.m., the midnight, rush hours in the morning, as well as rush hours in the evening. At midnights, some people tend to be fatigued or drunk in the restaurant when they are driving. In morning rush hours, our model captures the pattern of peak hours of accident occurrence explosion and predict a series of high-risk regions in the map and alarm citizens to raise their vigilance and take care to drive on the road. Generally, our model can foresee the near future and provide people with a road network risk map for reference.

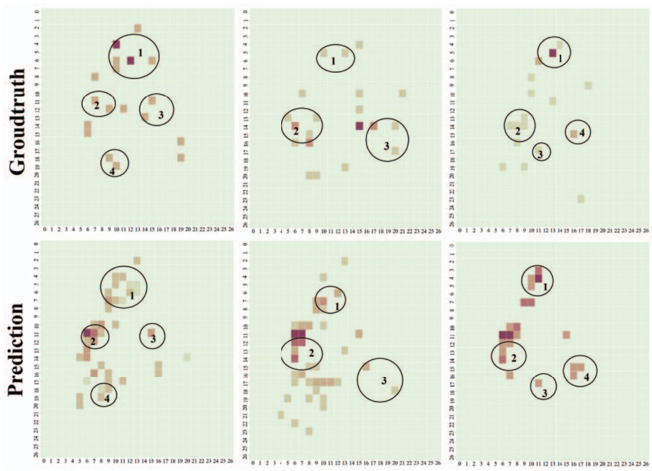

Fig. 4: Visualization of accident risk prediction

V. Conclusion and Discussion

In this paper, we propose a novel multi-task ResNet based framework which captures both spatial and temporal dependency to predict the hour-level citywide traffic accident risk. We exploit the correlations between strength of human mobility and citywide accident risk distribution, then incorporate the taxi trip records and speed into the accident risk forecasting task. Additionally, it is the first time to introduce multi-task learning to accident prediction problem and we reweight the different spatial and temporal dependency, enabling our model surpasses other alternative methods. Due to our model has nice scalabilities, thus it can be extended to other metropolitan areas and establish a real-time accidents risk warning system online. Both hour-level and long-term prediction model can be easily extended to help police force arrangements and citizens' safety travel plan. For future work, we should devise new methods to suppress the local relative low-risk regions and address the short-term accident sparsity problems. In terms of the limitations of our model, we still have a long way to promote the prediction accuracy. One feasible way is to incorporate other urban data such as smartphone records, information of individual vehicles. Also, more ITS data collection system can be deployed on the roads to provide more valuable information for the city brain.

References

[1] Q. Chen, X. Song, Z. Fan, T. Xia, H. Yamada and R. Shibasaki, "A Context-Aware Nonnegative Matrix Factorization Framework for Traffic Accident Risk Estimation via Heterogeneous Data," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 2018, pp. 346-351.

[2] C. Chen, X. Fan, C. Zheng, L. Xiao, M. Cheng and C. Wang, "SDCAE: Stack Denoising Convolutional Autoencoder Model for Accident Risk Prediction Via Traffic Big Data," 2018 6th International Conference on Advanced Cloud and Big Data (CBD), Lanzhou, GanSu, China, 2018, pp. 328-333.

[3] Z. Yuan, X. Zhou and T. Yang, " Hetero-Convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatiotemporal data," 2018 Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining(KDD), London, United Kingdom, 2018, pp. 984–992.

[4] J. Zhang, Y. Zheng and D. Qi. "Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction" 2016 AAAI Conference on Artificial Intelligence (AAAI), North America, February 2017, pp. 1655-1661.

[5] J. Jonathan, T. Peter, J. Timothy and K. Tapan, "Driver behavior during bicycle passing maneuvers in response to a Share the Road sign treatment," Accident Analysis & Prevention, vol. 70, pp. 92-99, September 2014.

[6] K. He, X Zhang, S. Ren and J. Sun. "Deep Residual Learning for Image Recognition," 2016 CVPR Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, North America, 2016.

[7] Z. Zhou. "Attention Based Stack ResNet for Citywide Traffic Accident Prediction," 2019 20th IEEE International Conference on Mobile Data Management (MDM). HongKong, China, 2019: 369-370.