2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)
October 8-12, 2022, Macau, China

Deep Learning-based Approach on Risk Estimation of Urban Traffic Accidents

Zhixiong Jin[1] Byeongjoon Noh[2] Haechan Cho[1] and Hwasoo Yeo[1]

2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC) | 978-1-6654-6880-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ITSC55140.2022.9922246

***Abstract*— This paper focuses on risk estimation problems of urban traffic accidents using deep learning approaches. There are two major challenges in the previous studies. The first challenge is the data imbalance problem that occurs numerous zeros in input data and can negatively affect the risk estimation results. The second challenge lies in neglecting the road environmental factors in risk estimation, which are also essential in causing traffic accidents. In order to address the aforementioned two problems, this study developed a hierarchical deep learning-based model with mobility and road environment data for estimating the risk of urban traffic accidents. The experiment results indicate the proposed method outperforms other existing models. The suggested method can be applied to the traffic warning system to assist people to avoid traffic accidents and further used in traffic accident prediction.**

I. INTRODUCTION

With the rapid development of urbanization, road safety becomes a major public health and social issue worldwide [1]. According to the World Health Organization (WHO), approximately 1.3 million people are killed and 50 million people are injured with disabilities every year, as a result of road traffic accidents [2]. Furthermore, the damage from traffic accidents has caused serious socioeconomic challenges in nations [3], [4]. In 2018, the Republic of Korea ranked high in terms of the number of traffic accidents occurred among OECD countries with approximately 200 thousand traffic accidents and 7.3 casualties per 100,000 people [5]. Therefore, it is necessary to establish strategies and develop systems in order to prevent traffic accidents.

In traffic accidents, the level of risk is governed by one or more factors such as human factors and road environmental factors as well as the external environment (e.g., weather conditions) [6]–[8]. Thus, it is imperative to analyze traffic accident data to investigate the relationships between risk outcomes and their related risk factors. In order to alleviate the adverse consequences of traffic accidents, it is also essential to develop models for inferencing the risk outcomes [9]. Most of the studies use injury severity or accident frequencies to reflect the level of the risk. The extensive studies have aimed at identifying the factors that significantly influence the injury severity of traffic accidents [2], [9]–[13]. For example, the authors in [9] proposed a framework for analyzing and predicting injury severity of traffic accidents based on factors contributing to the severity of collision injuries such as road types, weather, and light condition. They employed a stacked sparse autoencoder (SSAE) to consider the comprehensive factors for traffic accident analysis. The authors in [10] scrutinized the factors such as fatigues, genders, internal/external distractions (e.g., rushing to the destination, listening to the music, etc.), and investigated by surveying and simulating. The severity of injury used in these researches reflects the consequences of traffic accidents. Thus, researches on the injury severity can help decide the appropriate responses, such as enacting laws, repairing infrastructure, and deploying additional speed cameras.

Despite the fact that it is essential to determine the influencing factors of risk, proactive actions to prevent traffic accidents should be performed ahead of time. Some studies have focused on estimating and further predicting road traffic accidents and injury severity in urban by using mobility data as well as the contributing factors [14]–[20]. In general, the mobility data records information generated by activities, events or transactions [21]. For example, the authors in [14] proposed a traffic accident casualty prediction model using neural network and data mining techniques. They used historical data such as floating population, number of registered cars, and number of accidents. Some studies have adopted a concept of a hotspot where there is a high probability of an accident or traffic risk and estimated it in the spatial and temporal perspectives. The authors in [15] also proposed a traffic accident count prediction model using a Bayesian hierarchical approach. The proposed model can rank the candidate sites, called hotspots, according to their potential risk for future time period. Also the authors provided the simple diagnostics to validate the predictive capability of the proposed model. The authors in [20] computed traffic risk as well as traffic speed and flow with potential and broad usage of deep learning algorithms based on mobility data such as traffic data from infrastructure, trajectory data from vehicles, and Automatic Fare Collection (AFC) devices, which are widely deployed by urban transit systems (i.e., subway, buses, taxis).

Even if the previous studies have achieved significant success, there are two main challenges in estimating risk for urban traffic accidents. One challenge is the data imbalance problem that occurs when too many zeros are observed in the input data. Collecting a great number of data for model development is difficult due to the rarity of the accidents. Furthermore, it is complex to estimate the risk accurately since the input data is contaminated with zero

[1]Zhixiong Jin, Haechan Cho and Hwasoo Yeo with the Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea {iziz56, gkqkemwh, hwasoo} @kaist.ac.kr

[2]Byeongjoon Noh with the Applied Science Research Institute, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea powernoh@kaist.ac.kr

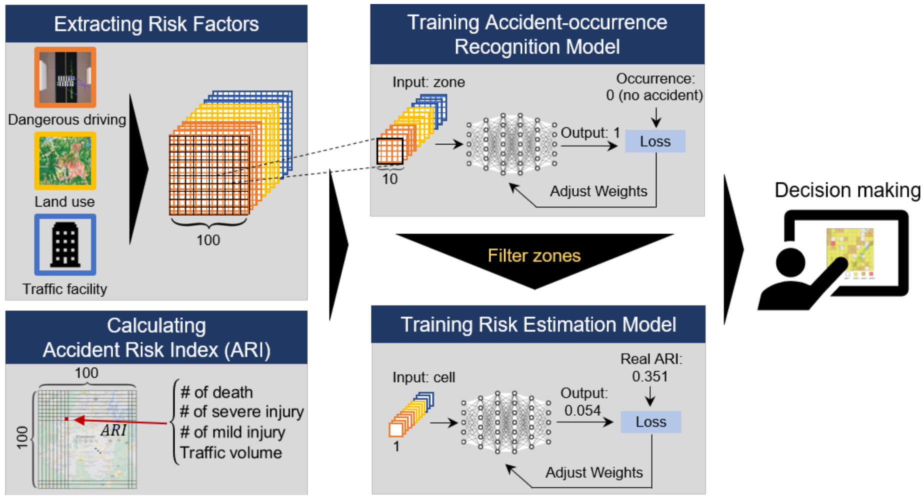

Fig. 1. The framework of risk estimation of urban traffic accidents.

values. The other challenge is related to data used in the model development. Most of the studies used mobility data as their primary input, however, other factors such as land use and traffic facilities can also be the triggers of the traffic accidents.

In order to address the aforementioned two problems, this study developed a hierarchical deep learning-based model for estimating the risk of urban traffic accidents. The proposed model is built using large-scaled mobility and road environment data. The model framework is shown in Fig 1. First, we discretized the target area into 10,000 cells and preprocessed the big traffic data for each cell. Then the hierarchical deep learning model consisting of Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs) was built based on the preprocessed data. The purpose of the proposed model is to reduce the number of zero values in various training data while also improving estimation performance.

II. Data Preprocessing and Description

For the risk estimation on the traffic accidents, we first need to quantify the level of the risk and then find out risk factors. Accordingly, we used the Accident Risk Index (ARI) as a surrogate measure of risk. For risk factors, we employed the aggressive driving and road environment factors since we considered that both of them can affect the estimation of risk in urban traffic accidents.

In our experiment, the study area is Daejeon city. We discretized the study area into 100 by 100 cells and set each cell in the grid as a basic spatial unit of the risk estimation (Fig 2). The width and height of each cell are approximately 230m.

A. Accident Risk Index (ARI)

Using the traffic accident counts alone to represent the risk does not take into account the differences between severe and slight injuries caused by traffic accidents. Also, traffic volume is one of the critical causal factors of traffic accidents. To overcome the limitation of using traffic accident counts as a measure of the risk, we proposed the ARI, which is defined as,

$$ARI = \frac{w_1 * DEATH + w_2 * SERI + w_3 * SLTWD}{\overline{V}} \tag{1}$$

where $DEATH$, $SERI$, and $SLTWD$ are the number of deaths, seriously injured, and slightly wounded injured cases, respectively. w_1, w_2, and w_3 are weights for the corresponding severity of injury. In this work, we used $w_1 = 1$, $w_2 = 0.7$, and $w_3 = 0.3$ based on the standards from Korea Transportation Safety Authority [22]. $\overline{V}$ represents the average traffic volume.

We computed the ARI for each cell on the grid in the study area using road accident and traffic volume data in 2019. A total of 8,337 traffic accidents occurred in the target area. Fig 3 (a) and Fig 3 (b) illustrate the grid-based representation of the number of traffic accidents and ARI in Daejeon City in 2019, respectively.

B. Aggressive Driving

The DTG (Digital Tachograph) data is mobility data that contains real-time vehicle information such as location, speed, and direction, collected by the onboard unit on the vehicle during the driving. We used DTG data collected by probe vehicles operated by local governments and commer-

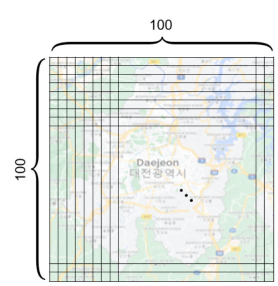

Fig. 2. Grid-based representation of the target area.

cial vehicles (e.g., bus, taxi, and truck) managed by the Korea Transportation Safety Authority in the Republic of Korea.

From DTG data, we identified eleven aggressive driving such as sudden start, sudden turn, and overspeed based on the rules defined by the Korea Transportation Safety Authority [22]. Fig 4 represents the process of extracting aggressive driving behaviors from the DTG data.

C. Road Environment

We also considered road environment as the risk factor since traffic accidents would occur due to the absence of a proper warning sign in infrastructures [6]. We used two types of road environment factors: land use and traffic facility.

Land use classified the land into residential, commercial, industrial, and green areas. Since land use determines the general direction of urban or suburban development, the overall traffic characteristics of the cells are highly dependent on land use and thus affect driver's behavior. For example, drivers tend to be more cautious around the residential area.

Traffic facilities including traffic lights, CCTV, as well as warning signs can be one of the factors that affect the accident risk in the urban area. To be more specific, proper placement of warning signs would reduce the number of traffic accidents. We calculate the number of each traffic facility to allow measuring the strengths of the respective preventive measure in the cell.

III. Deep Learning-based Risk Estimation

The core of the proposed model is to estimate the risk of each cell. To achieve it, we divided the whole process into two steps. The first step is to determine whether or not the accident occurred in the target region. Then in the second step, we estimated the risk of each cell based on the first step's result. The ARI represents the level of risk in this study. The input data consists of three types of data: aggressive driving (dri), land use (land) and traffic facility (fac) data. Fig 5 depicts the associated model design. The whole process can be written as follows,

$$\begin{aligned} ARI^i_{Cell} &= f_{model}(Cell_i) \\ Cell_i &= [dri_i, fac_i, land_i] \end{aligned} \tag{2}$$

where $Cell_i$ is the input data of i^{th} cell, f_{model} means the function of proposed model and ARI^i_{Cell} is the estimated ARI of the i_{th} cell.

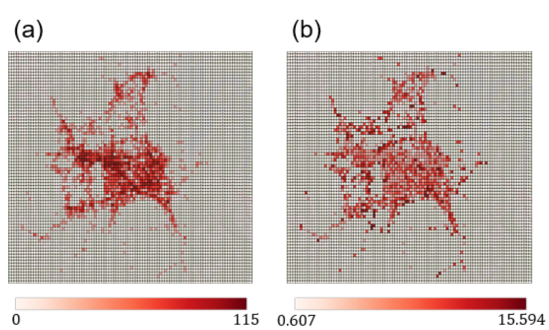

Fig. 3. (a) Traffic accident counts and (b) ARI in the target area.

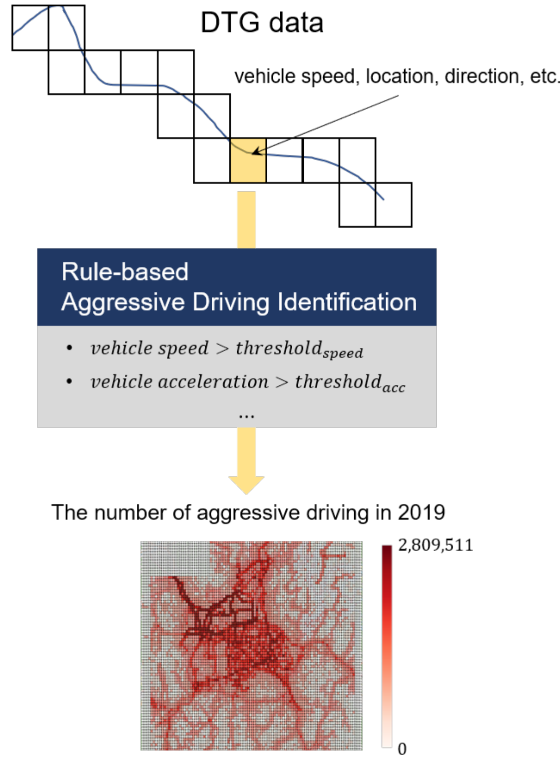

Fig. 4. Process of estimating the number of aggressive driving and representation in the target area in 2019.

A. Accident-occurrence Recognition Model

The main purpose of the model is to recognize the regions of accident occurred based on the input data. This method

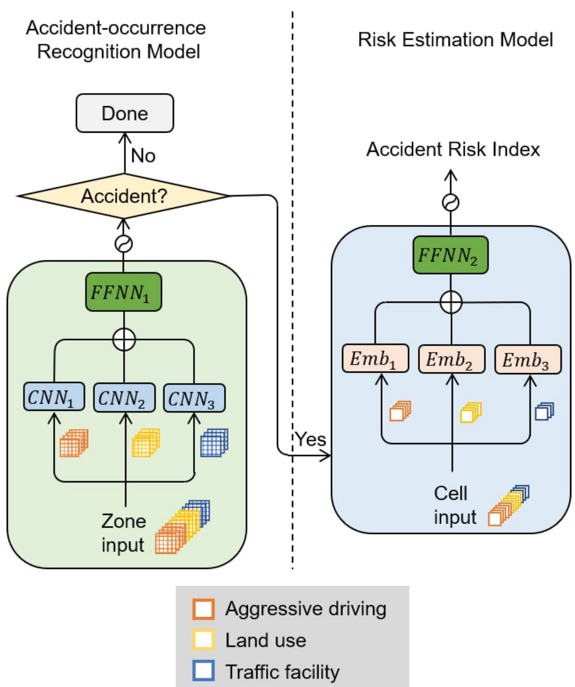

Fig. 5. Model architecture.

is efficient to deal with imbalanced accident data problems that excessive zero observations exist in input data. In other words, if we estimate ARI from the input directly, the final results will be highly influenced by a great number of zero values even if they show high accuracy. The results can be efficiently used in developing the ARI estimation model.

The model mainly consists of three Convolutional Neural Networks (CNNs) with a multi-layer Feed-Forward Neural Network (FFNN). It is demonstrated that CNN shows high performance in image features extraction. Similarly, the major roles of the three CNNs in this study are to extract the features of input data from the perspectives of aggressive driving, land use, and traffic facility. In this step, instead of using the input data directly, we clustered the cells in the target area by $N \times N$ cells, $N < 100$. The term "zone" is used to describe the newly clustered cells. The reason behind this is to reduce the effects of excessive zero values in CNN model training. In zone generation, we clustered the cells that are linked with each other to preserve the spatial properties of the target area. The generated zone consists of $(\frac{100}{N})^2$ cells. In this study, we defined $N = 10$. After extracting the useful information from three different input data, we used FFNN to build a classification model. The function of FFNN in this step is to estimate the occurrence of accidents at each zone, $i.e., 0 : no\ accident, 1 : accident$. The whole process can be written as follows,

$$\begin{aligned} Zone_j &= Cluster(Cell_{n \cdots m}) \\ Label(Zone_j) &= \lfloor FFNN(CNN(Zone_j)) \rceil \end{aligned} \tag{3}$$

where $Cell_{n \cdots m}$ represents the i^{th} to m^{th} cells, $Zone_j$ is the j^{th} zone, $Label$ is 0 in case of no accident and 1 in case of accident occurred, and $\lfloor \cdot \rceil$ means rounding to nearest integer.

B. Risk Estimation Model

In the second step, we used the accident-occurrence information to further estimate the risk of each cell. The ARI is used to quantify the level of risk. According to the results from the previous step, we filtered the non-accident zones in our dataset, and the rest are further used in risk estimation.

The major component of the risk estimation model is a FFNN with three different embedding layers. In machine learning, the embedding layers convert data from a low-dimensional space into high-dimensional vectors. The embedding function makes it easier to represent the features of large and sparse vectors. In this study, we used three embedding layers to represent the different features of input data since each one has different property and consequence on the final estimation results. After representing the input data features by using embedding layers, we concatenated and applied them to FFNN. In this step, FFNN is used to estimate the ARI of each input cell. The whole process can be written as follows,

$$\begin{aligned} Cell^i_{zone_j} &= f_{zone}(Zone_j) \\ Cell^i_{emb} &= Embedding(Cell^i_{zone_j}) \\ ARI^i_{Cell} &= FFNN(Cell^i_{emb}) \end{aligned} \tag{4}$$

where $Cell^i_{zone_j}$ represents the i^{th} cell in the j^{th} zone, f_{zone} means the translation function from zone to cell, and $Cell^i_{emb}$ denotes the embedded i^{th} cell.

IV. Experiments and Results

In this section, we evaluate the performance of our proposed model with several baseline models, such as Support Vector Regression (SVR), Logistic Regression (LR) and Naïve Bayes Classification (NBC). The whole process is done with Intel(R) Core(TM) i9-9900 CPU @ 3.10GHz, 16.0GB RAM, and NVIDIA GeForece RTX 2060.

A. Experimental Design

In this study, we chose the data from 2019-01-1 to 2019-12-31 to do the experiments. The dataset consists of 3,650,000 cells. In the first experiments, we clustered 100 cells as a zone and the total dataset is compressed to 36,500 zones. We randomly selected 70 % of the data for training the accident-occurrence recognition model and the rest data (30%) for testing. In the second experiment, we filtered the non-accident zones and separated the processed data in the same manner (70% for training and 30% for testing).

The architecture of the proposed model is built on Pytorch. In accident-occurrence recognition model training, the loss function is Binary Cross-Entropy (BCE), and the Adam optimizer is used via back-propagation of the model. In risk estimation model training, we chose Mean Square Error (MSE) as the loss function and Adam as the optimizer.

In this study, we added a dropout layer with a 0.2 dropout rate between each fully connected layer in FFNN to avoid overfitting. In addition, Rectified Linear Units (ReLU) and Sigmoid are used as activation functions during the model training.

B. Performance Evaluation

1) Evaluation Metrics: The model evaluation is divided into two steps. In the first step, the performance of the accident-occurrence recognition model was measured with its four metrics - accuracy, precision, recall and F-score. The F-score is one of the common methods in statistical analysis of binary classification. The traditional F-score uses the harmonic mean of precision and recall, which is defined as:

$$F_{score} = \frac{2}{\frac{1}{recall} + \frac{1}{precision}} = 2 \times \frac{precision \times recall}{precision + recall} \tag{5}$$

In the second step, to evaluate the risk estimation model, we selected Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) as our metrics, which are defined as:

$$\begin{aligned} MSE &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \\ RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \\ MAE &= \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \end{aligned} \tag{6}$$

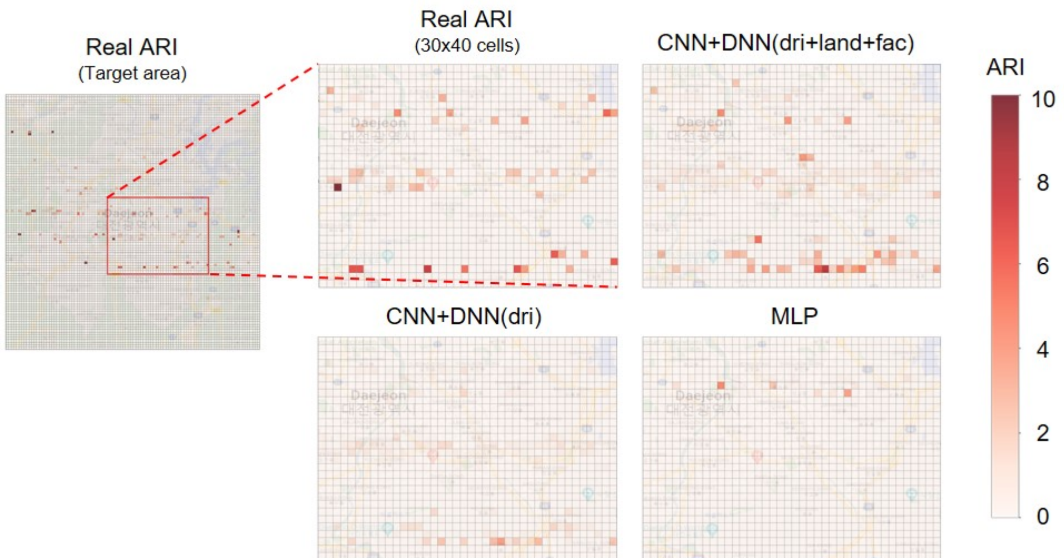

Fig. 6. Estimation results of CNN+DNN(dri+land+fac), CNN+DNN(dri), and MLP on the subset (30x40 cells) of the target area.

where n is the sample size, y_i and $\hat{y}_i$ are real and estimated ARI respectively.

2) Baseline Models: We chose the several traditional methods to compare the performance with our proposed risk estimation model. The baseline models are Support Vector Regression (SVR), Logistic Regression (LR), Naïve Bayes Classification (NBC), and Multilayer Perceptron (MLP). We also conducted an ablation study that compared the results with different input data to show the efficiency of land use and traffic facility data in an accident-occurrence recognition model.

3) Performance Evaluation: In this study, two experiments are implemented to evaluate our proposed models. In the first experiment, we compared the performance of the model with baselines on the test set, and the results are shown in Table I. The test results demonstrated that the CNN model that is trained with aggressive driving, land use, and traffic facility data outperforms baseline models. In addition, we can illustrate that it is efficient to use aforementioned data to improve the model performance.

After filtering the non-accident occurrence zones based on the proposed model, we estimated the ARI of each cell, and the results are shown in Table II. We used the MLP as the baseline model in this experiment. The results demonstrated that the proposed model, which is the combination of CNN and DNN, shows better performance than the baseline model. Similar to the first experiment, we also used various input data to indicate their efficiency on model performance improvement. The result showed that using three types of data is beneficial to improve the model performance. To evaluate the effectiveness of the proposed model, we selected the traffic accidents data during the first week of August in 2019 and depicted it in Fig 6. We observed that our model identified the high-risk regions effectively which the model without environmental factor (CNN+DNN(dri)) and non-hierarchical model (MLP) cannot detect.

TABLE I
PERFORMANCE OF ACCIDENT-OCCURRENCE RECOGNITION MODELS

Method	Accuracy	Precision	Recall	F-Score
SVR	0.903	0.673	0.292	0.407
LR	0.877	0.438	0.287	0.347
NBC	0.896	0.583	0.324	0.417
MLP	0.900	0.822	0.148	0.251
CNN (dri)	0.914	0.817	0.322	0.462
CNN (dri+fac)	0.930	0.793	0.573	0.665
CNN(dri+land)	0.929	0.698	**0.664**	0.681
CNN(dri+land+fac)	**0.935**	**0.848**	0.631	**0.724**

TABLE II
PERFORMANCE OF RISK ESTIMATION MODELS

Method	MSE	RMSE	MAE
MLP	0.173	0.415	0.103
CNN+DNN(dri)	0.171	0.413	0.102
CNN+DNN(dri+fac)	0.165	0.407	0.089
CNN+DNN(dri+land)	0.170	0.410	0.090
CNN+DNN(dri+land+fac)	**0.159**	**0.399**	**0.088**

V. CONCLUSION

In this study, we proposed a hierarchical deep learning-based model for estimating the risk of urban traffic accidents. The model was built based on three types of big traffic data - aggressive driving, land use, and traffic facility data. We also defined the ARI, a quantitative degree of traffic accident risk. We first used the CNN model to recognize and filtered the accident-occurrence region. Then DNN with different embedding layers was used to estimate the risk for each cell. The proposed method efficiently improves the model performance by reducing the effects of imbalanced accident data in model training and considering the different types of input data separately. The ablation studies revealed that using both aggressive driving, land use, and traffic facility factors was efficient to develop a high-performing risk estimation model. The performance comparison based on several evaluation metrics shows the accuracy and efficiency of our model. The proposed method is easy to implement in a traffic accident warning system, and it can assist people in avoiding traffic accidents by guiding them to safer regions. Also, this approach can be further used in traffic accidents prediction.

There are several directions in which the current study can be extended to improve the model performance. The first one is to solve the deficiency of our model structure. In our study, we used the CNN model to identify the accident-occurrence regions and filter out the rest. Even though the model shows high accuracy according to Table I, it does not perform well on other metrics. This outcome will result in more zero values in risk estimation model training, lowering its performance. Second, we require more various types of data in model development. Specifically, other variables, such as school zones, special events, and weather conditions, can affect the results in addition to the data included in this study. To the best of our knowledge, this approach is a first attempt to use a hierarchical model structure with mobility and road environment data to estimate the risk of urban traffic accidents. To further improve the model efficiency, we plan to solve the deficiency of model structure and incorporate the comprehensive factors in model development. We expect that future work with a developed model trained by various types of traffic data will be promising to make better estimation results. In addition, we will divide the target region into different levels such as intersections and links to further assess the model efficiency.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00784, Development of intelligent road and traffic situation detection technology using mobile infrastructure image sensor). We thank Korea Transportation Safety Authoriy for providing DTG data, and Daejeon City for providing safety facility data in Daejeon.

REFERENCES

[1] J. Yeo, J. Lee, J. Cho, D.-K. Kim, and K. Jang, "Effects of speed humps on vehicle speed and pedestrian crashes in south korea," *Journal of safety research*, vol. 75, pp. 78–86, 2020.

[2] L. Zheng, T. Sayed, and F. Mannering, "Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions," *Analytic methods in accident research*, vol. 29, p. 100142, 2021.

[3] K. Goniewicz, M. Goniewicz, W. Pawłowski, and P. Fiedor, "Road accident rates: strategies and programmes for improving road traffic safety," *European journal of trauma and emergency surgery*, vol. 42, no. 4, pp. 433–438, 2016.

[4] M. Ruikar *et al.*, "National statistics of road traffic accidents in india," *Journal of Orthopedics, Traumatology and Rehabilitation*, vol. 6, no. 1, p. 1, 2013.

[5] KoreaIndex, "Traffic accident status (death, injury)," https://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1614&param=003/, 2021, [Online; accessed 18-December-2021].

[6] P. Kopelias, F. Papadimitriou, K. Papandreou, and P. Prevedouros, "Urban freeway crash analysis: geometric, operational, and weather effects on crash number and severity," *Transportation Research Record*, vol. 2015, no. 1, pp. 123–131, 2007.

[7] J. De Oña, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury severity on spanish rural highways using bayesian networks," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 402–411, 2011.

[8] L.-Y. Chang and H.-W. Wang, "Analysis of traffic injury severity: An application of non-parametric classification tree techniques," *Accident Analysis & Prevention*, vol. 38, no. 5, pp. 1019–1027, 2006.

[9] Z. Ma, G. Mei, and S. Cuomo, "An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors," *Accident Analysis & Prevention*, vol. 160, p. 106322, 2021.

[10] G. Fountas, S. S. Pantangi, K. F. Hulme, and P. C. Anastasopoulos, "The effects of driver fatigue, gender, and distracted driving on perceived and observed aggressive driving behavior: A correlated grouped random parameters bivariate probit approach," *Analytic methods in accident research*, vol. 22, p. 100091, 2019.

[11] Q. Hou, X. Huo, J. Leng, and F. Mannering, "A note on out-of-sample prediction, marginal effects computations, and temporal testing with random parameters crash-injury severity models," *Analytic Methods in Accident Research*, vol. 33, p. 100191, 2022.

[12] A. Arun, M. M. Haque, S. Washington, T. Sayed, and F. Mannering, "A systematic review of traffic conflict-based safety measures with a focus on application context," *Analytic Methods in Accident Research*, vol. 32, p. 100185, 2021.

[13] N. Alnawmasi and F. Mannering, "The impact of higher speed limits on the frequency and severity of freeway crashes: Accounting for temporal shifts and unobserved heterogeneity," *Analytic Methods in Accident Research*, vol. 34, p. 100205, 2022.

[14] G. A. Ali and A. Tayfour, "Characteristics and prediction of traffic accident casualties in sudan using statistical modeling and artificial neural networks," *International journal of transportation science and technology*, vol. 1, no. 4, pp. 305–317, 2012.

[15] L. Fawcett, N. Thorpe, J. Matthews, and K. Kremer, "A novel bayesian hierarchical model for road safety hotspot prediction," *Accident Analysis & Prevention*, vol. 99, pp. 262–271, 2017.

[16] S. He, M. A. Sadeghi, S. Chawla, M. Alizadeh, H. Balakrishnan, and S. Madden, "Inferring high-resolution traffic accident risk maps based on satellite imagery and gps trajectories," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 11 977–11 985.

[17] H. Ren, Y. Song, J. Wang, Y. Hu, and J. Lei, "A deep learning approach to the citywide traffic accident risk prediction," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018, pp. 3346–3351.

[18] Z. Liu and K. Wu, *Mobility Data-Driven Urban Traffic Monitoring*. Springer, 2021.

[19] J. Bao, P. Liu, and S. V. Ukkusuri, "A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data," *Accident Analysis & Prevention*, vol. 122, pp. 239–254, 2019.

[20] Z. Liu, Z. Li, K. Wu, and M. Li, "Urban traffic prediction from mobility data using deep learning," *Ieee network*, vol. 32, no. 4, pp. 40–46, 2018.

[21] NACTO, "MANAGING MOBILITY DATA," https://nacto.org/wp-content/uploads/2019/05/NACTO_IMLA_Managing-Mobility-Data.pdf, 2019, [Online; accessed 18-December-2021].

[22] KTSA, "Dangerous driving," https://www.kotsa.or.kr/eng/engMain.do, 2021, [Online; accessed 10-November-2021].