Data-Driven Prediction for Red-Light-Running at a T-Junction

Sainan Zhang[1], Jun Zhang[1*], Weiguo Song[1], and Longnan Yang[1]

[1] University of Science and Technology of China, Hefei 230026, China
junz@ustc.edu.cn

Abstract. Running red lights is a serious road safety problem worldwide, which often leads to severe injuries and fatalities. Most recent works focus on identifying red-light-running behavior through surveillance cameras for punishment of violations. A few works predict the red-light-running behavior of drivers at intersections with Support Vector Machines (SVM) method. But they pay little attention to non-motor vehicles and the accuracy needs to be further improved. To address this problem, we conduct an observational experiment and construct a trajectory dataset (RedRun dataset) with the software Petrack. We also propose an Environment-Aware Red-light-running and Trajectory prediction Network (EA-RTN). It predicts the trajectories and red-light-running behavior of individuals (i.e. pedestrians, bicycles, electric vehicles, tricycles and cars) at T-junctions to help road users judge others' movement in advance. Specifically, EA-RTN consists of two modules: one is a fully connected neural network (FCNet), which uses two hidden layers to predict whether a road user will run a red light. The other is a two-layer long short-term memory neural network. It predicts the trajectories of road users in the next 2 seconds and then assists drivers to plan ahead. The losses of these two tasks are combined to update the weights for realizing the multi-task learning. To evaluate our model, experiments are conducted on RedRun dataset. The results show that our approach predicts red-light-running behavior of road users more accurately. The accuracy is about 10% higher than SVM method.

Keywords: Red-light-running, T-junction, Data-driven.

1 Introduction

The red-light-running behavior of road users is the main cause of traffic accidents, which seriously affects people's life and property safety. Although traffic lights have been set up at many intersections, and drivers who run red lights are punished through monitoring, they are mainly for motor vehicle violations. The behavior of non-motor vehicles running red lights is still not uncommon[1]. It is worth noting that predicting the trajectories of other road users and judging whether they will run a red light is very important and helpful in preventing traffic accidents.

Some scholars use deep learning methods to judge whether the vehicle runs a red light. They generally identify the vehicle's position and the interest area, like stop line,

K. R. Rao et al. (eds.), *Traffic and Granular Flow '22* , Lecture Notes
in Civil Engineering 443, https://doi.org/10.1007/978-981-99-7976-9_52

in the surveillance video[2]. If there is an intersection between these two positions, the vehicle will be considered to run a red light. However, they cannot make a prediction for red-light-running. A few scholars use support vector machine method[3] to make prediction, but the accuracy still needs to be improved. In this paper, firstly, we analyze the road users' behavioral characteristics in an observation experiment conducted at a T-junction. On this basis, we propose a data-driven model named EA-RTN to predict red-light-running behavior and the trajectories of road users.

Our model consists of two modules, as shown in Fig. 1, a two-layer fully connected neural network (FCNet) is proposed for predicting whether the road user will run a red light. The LSTM module is used for predicting the trajectories in the future two seconds with only the historical trajectories of 0.8 seconds as the input. To evaluate the performance of our approach, we conduct the experiment on the RedRun dataset. The results show that our model surpasses the Support Vector Machines (SVM) method and predicts the red-light-running behaviour more accurately. For the subtask of trajectory prediction, the average position error of each point for our method is 0.6 meter, which can assist drivers with planning ahead and avoid possible accidents.

The rest of this article is structured as follows: the observational experiment scenario and RedRun dataset construction are described in Section 2. Based on the analysis of the experiment, the EA-RTN model is proposed in Section 3. In Section 4, we evaluate our model EA-RTN on the RedRun dataset and compare the results with SVM. In Section 5, we summarize this work.

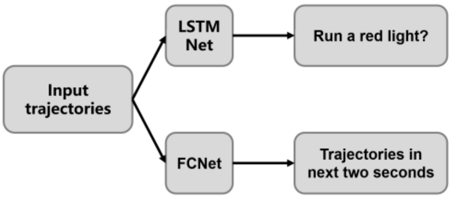

Fig. 1. The pipeline of our proposed RA-RTN.

2 Dataset

Due to the lack of trajectory data of road users at T-junctions, especially for crowded scenes such as school gates, we perform an observation experiment at the entrance of a university to analyze their motion characteristics. As shown in Fig. 2, the scenario is composed of a main road and a branch road, where the road widths are 15 meters and 10 meters, respectively. Traffic lights have been installed on the main road. The time interval when the traffic light is green, that is, the time for students to cross the main road, is between 20 seconds and 40 seconds. To avoid contingencies in the experimental phenomenon, we recorded 19 complete intervals in which the traffic light was green. There are 460 road users are observed and recorded in our observation. They include the Heterogeneous pedestrians (such as students, elders and children), bicycles, electric bicycles, electric tricycles, cars and buses. The individuals in the observation were not informed of the purpose of the experiment in advance.

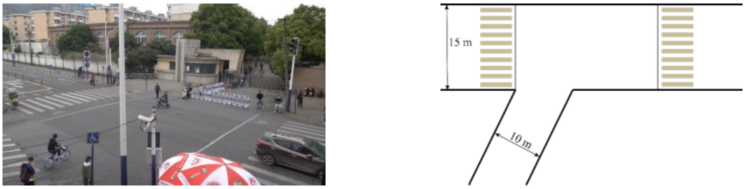

Fig. 2. The snapshot and schematic diagram of the observational experiment.

To reduce the distortion produced by the camera and get the motion information of each road user, we use the software Petrack to extract the experimental trajectories from our video recordings. To reduce the periodic interference caused by the head shaking, we use the average filtering method[4] to smooth the trajectory. The filter window size is set to 10 frames. Fig.3 shows the trajectories before and after filtering.

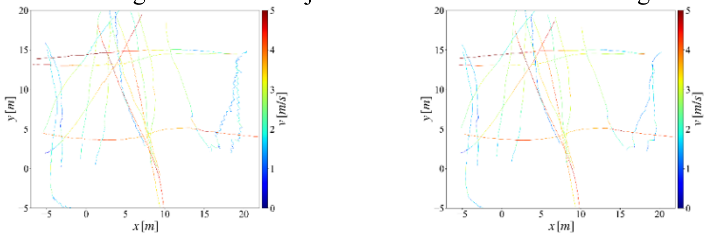

Fig. 3. Trajectories of road users. Left: raw trajectories extracted from the Petrack. Right: smooth trajectories obtained by the mean filter method.

In order to investigate the movement characteristics of red light runners, we counted the types and corresponding proportions of red light runners. As shown in Fig.4, 9.4% of road users have red-light-running behavior. Among them, 86% of red-light runners are electric bicycle users, 12% are electric tricycle users and 2% are pedestrians. This is the same as the findings of [5].

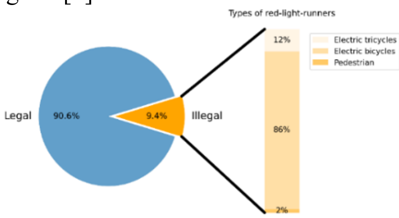

Fig. 4. A pie chart of the types of road users who run a red light.

To minimize the impact of trajectory error on red light running recognition, we select the position change over a period of time and the overall direction angle during this period for analysis. The time interval is set to 20 frames (0.8 seconds). We define that the y-axis is the same as the direction perpendicular to the main road. And the x-axis is the direction along the main road. As shown in Fig.5, these three features have different distribution on the illegal samples and legal samples. Here the illegal samples mean the road users who run red lights. Besides, we can observe that these three features of legal samples have more discrete values than those of illegal samples. The illegal samples have a larger offset in the x-axis direction than that in the y-axis. The heading angle between the time interval is measured with the feature $\frac{\Delta y}{\Delta x}$. For the illegal samples, the heading angle is closer to zero degrees.

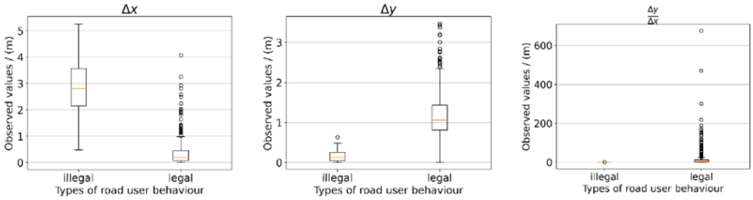

Fig. 5. Distribution of Δx, Δy, and $\frac{\Delta y}{\Delta x}$ for illegal and legal samples.

To verify and measure the difference between illegal and legal samples, we performed permutation tests for each feature. We construct the test statistic as the difference between the means of the illegal and legal samples for each feature. Our null hypothesis H_0 is that whether the sample is illegal or not has no effect on the distribution of features. The results are shown in Fig.6. The blue histogram is the distribution of the tests' results. The red dotted line is the result obtained in our observation. This shows that the values of these three features rarely occur under the null hypothesis H_0. The P values for each feature are 0.0, 0.0 and 0.037, which are all smaller than 0.05, indicating rejection of the null hypothesis H0.

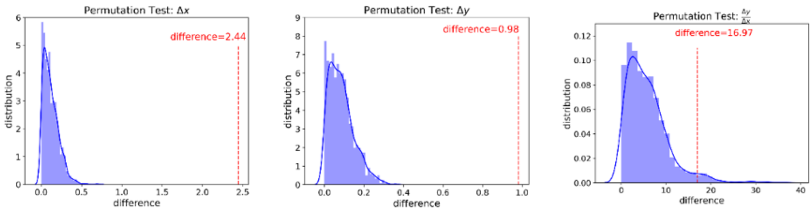

Fig. 6. The results of permutation test for Δx, Δy, and $\frac{\Delta y}{\Delta x}$.

3 Model

In this paper, we propose an environment-aware red-light-running and trajectory prediction network (EA-RTN). As shown in Fig.7, it consists of two modules: LSTM and FCNet. LSTM outputs the trajectories predicted in the future two seconds. FCNet outputs the probability of a road user running a red light. The combination of these two modules enables a multi-task learning.

3.1 LSTM

Motivated by the long short-term neural network (LSTM) which is capable of conveying and expressing information in long-term sequences effectively[6], we propose to use the combination of fully connected neural network (FC) and LSTM to learn the mapping from the historical trajectories to future trajectories. The input is the historical trajectories of 20 frames (0.8 seconds). In this module, the components of the trajectory on the x-axis and y-axis are learned separately. FC only has one layer. For the components on the x-axis, FCfirst extracts 64 features from the input for each moment. Then these features are put into a two-layer LSTM to learn the time dependence. After that, average pooling is performed to obtain a 128-dimensional feature vector, which is fed into a fully connected neural network to predict the position x in the next 50 frames (2 seconds). The same learning process is also used for the components on the y-axis. The predicted x and y coordinates are finally combined to get the trajectory in the future two seconds.

In this module, the input parameter is

$$\{x_i, y_{i,}\mathrm{i} = 1,2,\ldots,20\} \tag{1}$$

The output parameter is

$$\{x_j, y_{j,}\mathrm{j} = 1,2,\ldots,50\} \tag{2}$$

In order to calculate the error between predicted trajectories and the target values, we use mean squared error (MSELoss) as the loss function shown below

$$\mathrm{MSELoss(o,t)} = \frac{1}{n}\sum_{j=1}^{n}(\mathrm{t}-\mathrm{o})^2 \tag{3}$$

Where, n is the total number of training samples. o is the predicted position $(x_j, y_{j,})$ from LSTM. t is the corresponding target position.

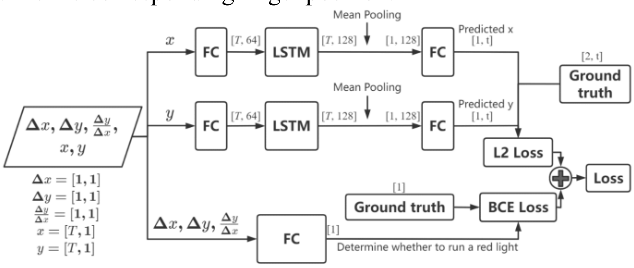

Fig. 7. The structure of EA-RTN. The plus sign indicates the feature fusion operation.

3.2 FCNet

Based on the analysis of the observational data, we found that the distributions of the features Δx, Δy, and $\frac{\Delta y}{\Delta x}$ are all statistically different on the illegal and legal samples. Therefore, it is possible to comprehensively judge the three features of the sample to predict whether it will run a red light. In this module, a fully connected neural network (FCNet) is used to achieve this goal.

The architecture of FCNet which has four layers is shown in Fig.8. The input parameters are

$$\{\Delta x, \Delta y, \frac{\Delta y}{\Delta x}\} \tag{4}$$

Where, for every sample, $\Delta x, \Delta y$ and $\frac{\Delta y}{\Delta x}$ all contain only one numeric value.

Whether the sample runs a red light is the target output. It only has two possible values: 0 or 1. The input layer, two hidden layers and output layer contain 3 neurons, 2 neurons and 1 neuron, respectively. In the FCNet, all hidden layers are followed by a rectified linear unit (ReLU) activation function. For the output layer, the sigmoid function is used to make its result between 0 and 1. If the output is larger than 0.5, the model predicts that the road user will run a red light.

In this module, we use binary cross entropy loss (BCELoss) as the loss function:

$$\mathrm{BCELoss}(p, t) = -\frac{1}{n}\sum_{i=1}^{n}(t_i \times \log(p_i) + (1 - t_i) \times \log(1 - p_i)) \tag{5}$$

Where, n is the total number of training samples. t_i is the target category of the ith sample. p_i is the predicted probability of the ith sample belonging to the target class.

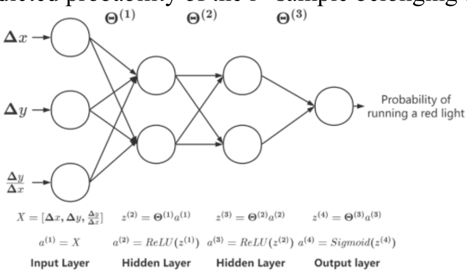

Fig. 8. The architecture of FCNet. Where X is the input vector. Θ is a set of parameters in the FCNet. z is the output of every layer. a is the result after activation.

3.3 Implementation details

To realize the multi-task learning of red-light-running prediction and trajectory prediction, we add up the losses of these two tasks and calculate it as the multi-task loss. The backpropagation algorithm is applied to train the EA-RTN. Adam is used to reduce the loss by optimizing the weights of our network. The learning rate is 0.001 at the beginning and then decays at the time of 0.999 every 25 epochs.

Considering the small sample size of the RedRun dataset, we adopt five-fold cross-validation method to train and verify our network. The samples are randomly divided into two parts: training-validation sets (account for 85%) and testing sets (account for 5%). These two datasets are independent with each other. In the experiments, we train a total of 500 epochs. The validation loss is used to choose the best epoch.

4 Experiments and results

To evaluate the performance of EA-RTN, we conduct the experiments on the RedRun dataset and compare the results of red-light-running prediction with support vector machine (SVM) method. The linear kernel is chosen as the kernel function of SVM. The numbers of training-validation samples and testing samples are 440 and 20, respectively. We define the completion of red-light-running behavior as the road user crossing the stop line opposite the intersection on the road when the traffic light is red.

4.1 Results

We visualize the confusion matrices of our model and SVM in Fig.9. We can observe that our model predicts the red light running individuals more accurately than SVM.

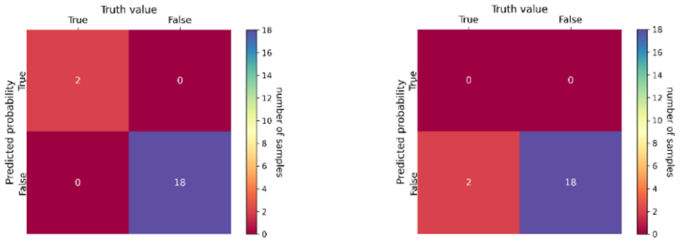

Fig. 9. The heatmap of confusion matrix for our model and SVM method. Left: EA-RTN. Right: SVM method.

We also calculate the accuracy and recall of the predictions in **Table 1**. The accuracy and recall for EA-RTN both reach 100%, which outperforms the results of SVM. We consider that this may be related to the simplicity of the scene and the small sample size, making it easy to learn the movement characteristics of red light runners. The performance of EA-RTN still needs to be evaluated on more datasets and scenarios.

Table 1. Quantitative Results of Network Performance Evaluation:

Model	Accuracy	Recall
SVM	90%	0%
EA-RTN	100%	100%

To quantitatively evaluate the performance of EA-RTN on the trajectory prediction task, we calculated the average value of the position error for each point on the trajectory. The formula is as follows:

$$position\ error = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2}\right) \tag{6}$$

Where, x_i , y_i , $\hat{x}_i$ and $\hat{y}_i$ are coordinates of each point on the true trajectories and the predicted trajectories. N is the number of test samples. n is the number of points on each trajectory.

On the test set, the error of EA-RTN is 0.6 m, which means that our model predicts the trajectories accurately and can help drivers make decisions in advance.

5 Conclusion

In this paper, we propose an approach EA-RTN to predict the red-light-running behavior and trajectories in future two seconds of road user in T-junction. LSTM module is used to learn the mapping between historical trajectories of 0.8 seconds and future trajectories of 2 seconds. FCNet is used to evaluate whether a road user will run a red light by taking into account the offset of the position in 0.8 seconds and the overall heading angle. Through the experiments conducted on the RedRun dataset, the performance of our model is evaluated. Our model outperforms the SVM method on the red light running prediction task. It is suitable for red-light-running prediction for T-junctions located at school gates, but it still needs to be evaluated on more datasets and scenarios.

6 Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 72174189， U1933105) and the Fundamental Research Funds for the Central Universities (Grant Nos. WK2320000043， WK2320000050).

References

1. Wu, C., Yao, L., Zhang, K.: The red-light running behavior of electric bike riders and cyclists at urban intersections in China: an observational study. Accident Analysis & Prevention. 49, 186–192 (2012). https://doi.org/10.1016/j.aap.2011.06.001.
2. Franklin, R.J.: Traffic signal violation detection using artificial intelligence and deep learning. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES). pp. 839–844. IEEE (2020).
3. Chen, X., Zhou, L., Li, L.: Bayesian network for red-light-running prediction at signalized intersections. Journal of Intelligent Transportation Systems. 23, 120–132 (2019). https://doi.org/10.1080/15472450.2018.1486192.
4. Kalman, R.E.: A new approach to linear filtering and prediction problems. (1960). https://doi.org/10.1115/1.3662552.
5. Qin, H., Wei, Y., Zhang, Q., Ma, L.: An observational study on the risk behaviors of electric bicycle riders performing meal delivery at urban intersections in China. Transportation research part F: traffic psychology and behaviour. 79, 107–117 (2021). https://doi.org/10.1016/j.trf.2021.04.010.
6. Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., Savarese, S.: Social lstm: Human trajectory prediction in crowded spaces. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 961–971 (2016).