CONCLUSION:

In this paper, six machine learning algorithms are applied to predict the occurrence of crime hotspots in a town in the southeast coastal city of China. The following conclusions are drawn:1) The prediction accuracies of LSTM model are better than those of the other models. It can better extract the pattern and regularity from historical crime data. 2) The addition of urban built environment covariates further improves the prediction accuracies of the LSTM model. The prediction results are better than those of the original model using historical crime data alone. Our models have improved prediction accuracies, compared with other models. In empirical research on the prediction of crime hotspots, Rummens et al. used historical crime data at a grid unit scale of 200 m×200 m, using three models of logistic regression, neural network, and the combination of logistic regression and neural network [41]. In the biweekly forecast, the highest case hit rate for the two-robbery type is 31.97%, and the highest grid hit rate is 32.95%; Liu et al. Used the random forest model to predict the hot spots in multiple experiments in two weeks under the research scale of 150 m × 150 m [23]. The average case hit rate of the model was 52.3%, and the average grid hit rate was 46.6%. The case hit rate of the LSTM model used in this paper was 59.9%, and the average grid hit rate was 57.6%, which was improved compared with the previous research results, For the future research, there are still some aspects to be improved. The first is the temporal resolution of the prediction. Felson et al. revealed that the crime level changes with time [43] Some studies have shown that it is useful to check the variation of risks during the day [44]. We chose two weeks as the prediction window. It does not capture the impact of crime changes within a week, let alone the change within a day. The sparsity of data makes the prediction of crime event difficult if the prediction window is narrowed down to day of a week or hour within a day. There is no viable solution to this challenging problem at this time. The second is the spatial resolution of the grid. In this paper, the grid size is 150m ∗ 150m. Future research will assess the impact of changing grid sizes on prediction accuracy. Third, the robustness and generality of the findings of this paper needs to be tested in other study areas. Nonetheless, the findings of this research have proven to be useful in a recent hotspot crime prevention experiment by the local police department at the study size.