1. **PRINCIPLES OF THEORETICAL CRIMINOLOGY IN PREDICTION OF CRIME HOTSPOTS:**

The focus of crime hotspot prediction is to forecast future concentration of criminal events in a geographical space. Theoretical criminology provides the necessary theoretical basis. Specifically, several related criminological theories not only provide guidance for us to understand the important influence of location factors in the formation and aggregation of criminal events, but also provide a basic mechanism for the police to use information of crime hot spots for crime prevention or control. It mainly includes routine activity theory, rational choice theory, and crime patterns theory. These three theories are generally considered as the theoretical basis of situational crime prevention. Routine activity theory [30] was jointly proposed by Cohen and Felson in 1979, and has now been further developed through integration with other theories. This theory believes that the occurrence of most crimes, especially predatory crimes, needs the convergence of the three elements including motivated offenders, suitable targets, and lack of ability to defend in time and space. Rational choice theory [31] was proposed by Cornish and Clarke. The theory holds that the offender’s choices in terms of location, goals, methods be explained by the rational balance of effort, risk and reward. Crime pattern theory [32] integrates the routine activities theory and the rational choice theory, which more closely explains the spatial distribution of criminal events. People form ‘‘cognitive map’’ and ‘‘activity space’’ through daily activities. At the same time, potential offenders also need to use their cognitive maps and choose specific locations for crimes in a relatively familiar space. When committing a crime, the offender tends to avoid those places they don’t know but to choose the places where the ‘‘criminal opportunity overlaps with cognitive space’’ based on their rational ability. The reason why these places become crime hotspots is that they have the obvious characteristics of ‘‘producing’’ or ‘‘attracting’’ crime. Therefore, the environmental factors of the places need to be considered besides historical crime data for the prediction of crime hotspots.

1. **BUILT ENVIRONMENT DATA :**

At present, a large number of studies show that the urban built environment has a significant impact on urban criminal behavior, through the impact of crime opportunities to reduce and prevent crime. In the 2007 Global Habitat Report, it was pointed out that the elements of the built environment have an important impact on the occurrence of criminal acts [33]. Point of interests (POIs) data and road network density data are considered as covariates in the crime prediction model.

1. POI DATA :

The urban infrastructure data POI includes the location information and attribute information of various urban facilities [34], [35]. Catering facilities, shopping malls and stores are usually located in places with convenient transportation and large flow of people, gathering a large number of different groups of people to generate the targets for the criminals, while entertainment places attract criminals [36]. These POIs are selected as covariates of the prediction model.

1. ROAD NETWORK DENSITY:

The conventional definition of road network density refers to total length of roads divided by the size of an areal unit. The area with a denser road network attracts greater flow of people, including potential victims and criminals. Previous studies have shown that the density of road network has an impact on crime rate, especially in public space [37].

1. **CRIME PREDICTION WITH MACHINE LEARNING ALGORITHMS:**

The traditional methods usually detect the crime hotspot area from the historical distribution of crime cases, and assume that the past pattern is to be repeated in the future [7], [2]. This assumption tends to be reasonable for predicting long-term stable crime hotspots. The commonly used KDE method can effectively identify such stable hotspot areas [10], [11]. The KDE method based on temporal autocorrelation tends to outperform the general KDE method [38] Liu et al. Compared the random forest and spatiotemporal KDE method, found that the random forest algorithm is more efficient than the traditional spatiotemporal KDE method in the smaller time scale and grid space unit [12] Gabriel et al. used the Gated Localized Diffusion Network for crime prediction at the street segment level [39]. Compared with the traditional Network-time KDE method, the diffusion network approach significantly increased the prediction accuracy. The ability of machine learning algorithm in processing non-linear relational data has been confirmed in many fields, including crime prediction. It has a faster training speed, can handle very high-dimensional data, and can also extract the characteristics of the data.