Cities 143 (2023) 104575

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Extracting the pickpocketing information implied in the built environment by treating it as the anomalies

Yao Yao [a,b], Anning Dong [a], Zhiqian Liu [a], Ying Jiang [a], Zijin Guo [a], Junyi Cheng [c], Qingfeng Guan [a], Peng Luo [d,*]

[a] School of Geography and Information Engineering, China University of Geosciences, Wuhan 430078, Hubei Province, China
[b] Center for Spatial Information Science, The University of Tokyo, Kashiwa-shi, Chiba 277-8568, Japan
[c] Institute of Remote Sensing and Geographic Information Systems, Peking University, 5 Summer Palace Road, Beijing 100871, China
[d] Chair of Cartography and Visual Analytics, Technical University of Munich, 80333 Munich, Germany

ARTICLE INFO

Keywords:
Pickpocketing crime
Street view
Deep anomaly detection
Interpretable analysis
Social disorder

ABSTRACT

The practice of crime risk mapping, enabled by the utilization of geospatial big data such as street view images, has received significant research attention. However, in situations where available data is scarce, mapping models may suffer from underfitting and generate inaccurate spatial pattern estimations of crime risk. The covert nature of pickpocketing crimes results in limited observed areas relevant to such criminal events, leading to insufficient coverage of geospatial data. Moreover, the location of crime is also influenced by socio-economic characteristics that may introduce biases into crime risk estimates. These factors render it challenging for the model to capture a valid crime risk pattern, potentially yielding misleading conclusions. Therefore, effectively extracting crime risk with limited data remains a challenge, especially when relying on easily accessible, widespread, and unbiased geospatial data. To address this challenge, we propose a novel crime risk assessment framework based on deep anomaly detection techniques, assuming that urban landscape anomalies carry deep crime risk information. We take Shenzhen as the study area and map the distribution of pickpocketing risk using street view images, accurately revealing the spatial aggregation of pickpocketing crime risk. Our findings indicate that pickpocketing crime in China is caused by regional economic conditions, built environment factors, and human routine activities. This study provides valuable insights for policing and prevention strategies aimed at addressing pickpocketing crimes in large Chinese cities. By leveraging our proposed crime risk assessment framework, decision-makers can allocate resources more efficiently and develop targeted interventions to mitigate crime risks.

1. Introduction

Crime has a significant impact on economic growth and human lives (ToppiReddy et al., 2018), a problem that has long plagued human societies. One of the most common crimes is Pickpocketing, which involves stealing a victim's property in a public or semi-public place (Deshotels, 2013). Pickpocketing is characterized by high concealment, small amounts of money involved, and high significant financial and material resources investment in detection and apprehension (Lafree & Birkbeck, 2010). Therefore, preventing pickpocketing yields greater policing benefits than detecting and apprehending the offender. The social disorder theory and crime pattern theory suggest out that an objective environment can stimulate crime generation to some extent (Shaw et al., 1942). Since pickpocketing requires physical contact with the victim and often occurs in urban environments, exploring the association between the urban environment and pickpocketing is vital for police departments to prevent and control this type of crime and maintain social stability.

Several previous studies have assessed crime risk using historical case data, spatio-temporal environmental data, or behavioral trajectory data. However, these approaches present certain limitations. Historical case data-based assessments can only consider crime patterns based on actual case occurrences and have low dimensions and a single source of information, making it challenging to assess crime risks in areas where

* Corresponding author.
E-mail addresses: yaoy@cug.edu.cn (Y. Yao), donganning@cug.edu.cn (A. Dong), liuzhiqian@cug.edu.cn (Z. Liu), snotra@cug.edu.cn (Y. Jiang), gzj2017@cug.edu.cn (Z. Guo), junyicheng@pku.edu.cn (J. Cheng), guanqf@cug.edu.cn (Q. Guan), peng.luo@tum.de (P. Luo).

https://doi.org/10.1016/j.cities.2023.104575
Received 20 May 2022; Received in revised form 18 August 2023; Accepted 21 September 2023

no crime has occurred or where crime data are unavailable (Hossain et al., 2020; Hu et al., 2018). Some researchers have integrated spatio-temporal environmental data (Ding & Zhai, 2021; Giménez-Santana et al., 2018) to further consider the spatio-temporal effects of the background environment on crime generation and evolution. Others have used micro-level behavioral trajectory data (Rumi et al., 2019; Xiao & Zhou, 2020) to assess crime risk, incorporating socioeconomic data such as demographic, GDP, and unemployment rates and trajectory data such like location check-in and cab traffic for crime risk assessment. However, obtaining fine-scale residential travel and socioeconomic data can be challenging in some areas, and, some environmental data may also be difficult to collect, leading to limited applicability and generalizability of existing methods. Thus, addressing the question of how to use easily accessible and equally objective indicators that can be correlated and mapped to crime instead of hard-to-obtain economic indicators remains a challenge.

In the field of crime risk mapping, street view imagery offers a potential solution due to its extensive coverage. Recent studies have shown that street view images can provide insight into the physical urban environment and reveal crime risk (He et al., 2017; Zhang et al., 2021). Street view imagery accurately depicts the physical urban environment and allows for inferences about urban perception (Wang et al., 2019a; Yao et al., 2019). With easy availability, high-frequency updates, and microscopic perspectives on the city, street view imagery has become an increasingly popular tool for analyzing human or physical environmental elements using semantic segmentation or target recognition methods to test crime theories (He et al., 2017; Yue et al., 2022). Zhang et al. (2021) recently analyzed Houston street view imagery and historical criminal records and found a discrepancy between people's perception of safety in the urban environment and the actual crime rate. However, most previous studies require large amounts of real, tagged crime data for analysis, which can be challenging to obtain for sparsely located crime events such as pickpocketing, whose data may have biased spatial distribution. Therefore, it remains unclear whether the relationship between street view imagery and crime can be effectively mined when dealing with sparse and biased data.

Pickpocketing is a very typical and common type of crime that affects people's daily lives, yet reliable data regarding its occurrence is scarce and biased. The available data may not accurately reveal the true spatial patterns of crime risk due to several factors. Firstly, crime data is scarce and incomplete in certain regions. For instance, in China, publicly available crime data primarily consists of judgment documents, which do not always reflect the true number of pickpocketing crimes committed. Due to the relatively minor nature of pickpocketing offenses, suspects often employ various methods to evade surveillance and avoid detection, resulting in underreporting of such crimes. Secondly, there is a problem of biased sampling in the available data. The spatial distribution of crime locations in judgment documents is not solely determined by the risk of crime but can also be influenced by population density, law enforcement efforts, economic conditions, and other attributes. As an example, densely populated and more economically developed areas have a high density of crime points, while sparsely populated suburban or rural areas may have limited data on pickpocketing, despite not necessarily having lower crime risks. Moreover, obtaining a conviction for a pickpocketing offense involves a complex process that includes the occurrence of the crime, police investigation, and court proceedings. Therefore, although judgment documents can serve as a reference for analyzing crime patterns, they may not fully capture the real spatial pattern of pickpocketing crime.

In conclusion, we contend that publicly available crime data, especially for pickpocketing, does not provide a comprehensive representation of the true spatial pattern of crime risk. Firstly, such data is highly sparse in space, which limits its utility in producing accurate crime risk assessments. Secondly, the pattern of crime data suffers from sampling bias, and can be influenced by socio-economic factors beyond crime risk considerations. While many studies have utilized multiple data sources to analyze crime risk, these studies often require large quantities of labeled data for training models (Hajela et al., 2021; Xiao & Zhou, 2020). However, given the limited amount of labeled crime data and the significant bias present at crime points, there are currently few effective methods for achieving accurate spatial predictions of global crime risk. Despite studies indicating that the collection of crime information by law enforcement agencies inevitably suffers from biases due to influences from the agencies themselves and those reporting the crimes, it is important to note that these data sources still exhibit fewer random biases compared to other sources, such as spontaneously reported crime victim survey data. Moreover, they provide accurate records of crime locations and processes, thus remaining a more trustworthy source of crime data (Brunton-Smith et al., 2023; Buil-Gil et al., 2022). However, when crime data is accurate but scarce, it remains unclear to what degree policing levels, as quantified by crime data such as judgment documents, can be trusted as reliable indicators of crime risk. Given the current limitations associated with relying solely on real crime points for analysis and decision support, we propose the following research question: How can precise predictions of global crime risk be generated when crime data are sparsely sampled and biased? If it proves possible to accurately extract crime risk information from such sparse data sources, particularly in relation to hidden crimes like pickpocketing, this could have significant implications for large-scale crime risk assessment and urban governance.

Due to various factors, it is common to conduct research on data with bias in the field of geographic information. Whether it's bias brought about in the data collection process (Li et al., 2016; Zhang & Zhu, 2019a),bias in geographically large data voluntarily uploaded by the public (Zhang, 2022; Zhang & Zhu, 2019b), or even bias in data collected by government agencies (Brunton-Smith et al., 2023; Buil-Gil et al., 2022), there are inevitable deviations. Although the data may be geographically biased, it is still numerically correct. We can trust that the more similar the geographical configuration (i.e., spatial neighborhood geographical variables) of two points (regions), the more similar the values (processes) of the target variable at these two points will be (Zhu et al., 2018). Based on this idea, finding suitable environmental features for the data and designing analysis methods that adapt to this data has become the key task in using biased data for geographical modeling.

This study proposes the Crime Anomaly Detection based on Street View (CADSV) framework, which utilizes deep learning methods to tackle the aforementioned challenges. Anomaly detection is a popular technique for identifying rare or unusual patterns within large datasets (Chandola et al., 2009), which is similar to a crime assessment task that extracts risk information from limited crime labeled street view images. In this study, we focus on the city of Shenzhen where we assess the risk of pickpocketing at various locations using judgment documents as to the supporting data. To further investigate the socioeconomic factors associated with crime, we incorporate point of interest (POI) data is used to represent the urban functional structures. Additionally, the random forest and SHapley Additive exPlanations (SHAP) techniques are used to utilize the complex relationship between the urban socioeconomic structure and spatial environment.

2. Related work

2.1. Risk assessment of pickpocketing crime based on spatial analysis

Crime risk assessment is essential in policing (Fan et al., 2021; Oswald et al., 2018). Some scholars have conducted crime risk assessments based on analysis of historical case data (Hossain et al., 2020; Hu et al., 2018). For instance, Hu et al. (2018) utilized a spatiotemporal kernel density estimation (STKDE) method to analyze the history of crimes in a particular location and identify burglary hotspots in the region. Similarly, Hossain et al. (2020) employed decision trees and the k-nearest neighbors (KNN) algorithm to evaluate crime risk using San

Francisco's criminal activity data from 2003 to 2015. However, these studies do not account for the interaction between crime and other social environment factors. The data used in these studies are typically low-dimensional and obtained from single sources, rendering them suitable only for macro trend statistical analyses with limited explanatory power for crime risk assessment. Furthermore, reliance on historical crime data from a specific location may limit the transferability of findings beyond that context.

The Broken Window Theory (BWT) elucidates the relationship between crime and environment, positing that visible signs of disorder and neglect can foster further criminal activity, including serious crimes (Wilson & Kelling, 1982). Certain scholars have augmented historical case data with spatiotemporal environmental data to better consider the spatiotemporal effects of the contextual environment on the generation and evolution of crime (Ding & Zhai, 2021; Giménez-Santana et al., 2018). For example, Giménez-Santana et al. (2018) used a risk-topography modeling approach to identify environmental factors associated with three types of violent crime events (homicide, assault, and theft) and assessed the risk for different crime types. Ding and Zhai (2021) used crime statistics and observed climate records in Beijing to demonstrate strong correlations between PM2.5, the Air Quality Index (AQI), and bus pickpocketing crimes. Based on these findings, they utilized a support vector machine approach was used to predict the risk of bus pickpocketing crimes. Many studies have demonstrated that crime generally tends to concentrate in micro-specific locations such as streets, thereby highlighting the importance of assessing crime risk at the micro-level for effective crime prevention and police control (Groff et al., 2010; Weisburd et al., 2004). However, spatio-temporal environmental data are often collected at the grid scale, which has limited spatial resolution, and is generally only suitable for macro-level studies while being insufficiently assessed at the micro-scale.

Crime pattern theory suggests that offenders do not randomly search for potential targets but instead rely on the path or routes of their daily activities to find suitable targets (Bernasco et al., 2013; Bernasco et al., 2017; Brantingham & Brantingham, 2013). Therefore, some scholars have integrated suspects' behavioral trajectory data with spatio-temporal environmental data (Bouma et al., 2014; Rumi et al., 2019; Xiao & Zhou, 2020). Notably, Zhao and Tang (2017) employed POI check-in data, weather data, and public service complaint data to predict future crimes in New York City. Results showed that the inclusion of dynamic data characterizing daily human activity helped to accurately assess crime risks. Hajela et al. (2021), meanwhile, constructed distinct crime prediction models using taxi data, historical crime data, and demographic data, comparing their effectiveness against each other. This study demonstrated that methods incorporating dynamic data are more effective in crime prediction than those relying exclusively on crime data or data pertaining to social environmental factors. In summary, these studies consider the impact of the environment on crime at a finer scale, which is more effective in assessing pickpocketing risk at the micro-scale. However, such research typically requires low-accessibility data, thereby limiting its applicability to larger areas. Conversely, street view data can satisfy both environmental information provision and large-scale information provision, providing data support for crime risk assessment.

2.2. Street view image and city perception

Street View Images (SVI) are composed of panoramic images of various locations on the street that provide a comprehensive reflection of the physical urban environment and human activities on a large scale (Kang et al., 2020; Yao et al., 2019). In comparison to behavioral trajectory data, SVIs are low-cost and highly accessible. Additionally, they can capture detailed information in the physical environment more comprehensively using a perspective similar to that of the human eye (Zhang et al., 2020). As such, they have been integrated into diverse urban studies, including urban safety perceptions (Wang et al., 2019b; Zhang et al., 2021) and urban crime research(He et al., 2017). For instance, He et al. (2017) employed used Google Street View to identify factors in the physical environment of Columbus cities that contribute to violent criminal activity. Results showed positive associations between crime rates and street graffiti, abandoned buildings, and abandoned cars. Similarly, Zhang et al. (2021) analyzed street view images and historical criminal records in Houston, finding that areas where people feel unsafe do not correlate with high crime rates. There existed a perceived bias between perceived safety and actual crime rates in the urban environment.

Most of the current studies examining the relationship between street view images and crime risk have focused on Western cities. However, it remains uncertain questionable whether research findings on Western cities can be effectively applied to Chinese cities. Firstly, there are significant disparities in architectural and urban planning styles between the East and West (Ashihara & Riggs, 1983). Secondly, various socio-political factors contribute to the differences in crime patterns between East and West (Farrell & Bouloukos, 2001; Steffensmeier et al., 2017). For instance, a comparative study of high school students in China and the United States revealed that crime rates were significantly lower in Chinese schools were much lower than in American ones (Webb et al., 2011). As such, it is crucial to investigate the relationship between urban environments and crime risk in China using street view images.

Regarding the use of street views for crime risk prediction, a typical approach involves first extracting high-dimensional semantic features from the images, followed by constructing regression models to establish the relationship between these features and crime risk. For example, semantic segmentation can be used to extract the proportion of green space within an image, while target detection can estimate the number of people present (Hipp et al., 2021; Jing et al., 2021). Street view images contain vast amounts of semantic information that humans have yet to explicitly express. This implicit information has the potential to uncover crime risk. However, current solely rely on human-defined semantic features, thus overlooking a large amount of semantic information present in the images. Moreover, the above framework encounters difficulties when analyzing risks in areas not covered by LBS data, particularly with biased and spatially sparse crime data. To address this problem, it is crucial to extract crime risks from street views based on sparse data, which would fill this gap and facilitate the building of end-to-end models by eliminating complex image processing steps.

2.3. Geographical research based on biased data

In fact, in the field of geographic information, analyzing data with a geographical distribution bias is a widely studied problem. For example, Volunteer Geographic Information (VGI), voluntarily uploaded by citizens, is one of the newly emerged types of big geographic data in recent years (Zhang & Zhu, 2018). With its rich geographic information, high update frequency, and low cost, VGI is used to reveal spatiotemporal patterns of geographic phenomena. However, since the spatial distribution of volunteer observation work is neither random nor regular, the observation results are often spatially biased towards areas with high population density or high route accessibility, leading to bias in geographic distribution. To address this issue, researchers typically start by comparing sample locations with the environmental covariates of the predicted areas, improving sample representativeness by comparing similarities. Based on this concept, studies by (Zhang & Zhu, 2019b) and (Zhang, 2022) have mitigated the spatial bias in samples based on VGI by improving sample representativeness.

Additionally, studies based on geographically biased data in the field of digital soil mapping underscore the importance of environmental covariates. Since the spatial distribution of soil samples may lean towards specific geographical areas and be influenced by the personnel taking measurements, soil samples are a type of data that can be easily affected by spatial bias (Li et al., 2016; Zhang & Zhu, 2019a). To solve this problem, (Fan et al., 2020) proposed the SoLIM-FilterNA method, which predicts the soil property values of unknown units by learning the characteristics of one or more environmental covariates, as long as the uncertainty of that unit does not exceed the set threshold. (Zhu et al., 2018) proposed a method that does not use the explicit relationships derived from the entire sample set, but instead makes predictions based on the comparison of the geographical configurations of the sample points and the prediction points. This study suggests that accurate spatial predictions can be made based on biased samples. Similarly, facing the issue of data sparsity, (Du et al., 2020) proposed a semi-supervised machine learning method for predictive mapping, which uses the natural aggregation (clustering) pattern of environmental covariate data to supplement the limited samples in prediction. The characteristic of these studies is that they step outside of the spatial dimension to deal with spatially biased data. It can be seen that in the case of geographical bias, if suitable environmental features and methods can be selected to fit the data, results better than traditional geographic models can be achieved.

2.4. Deep anomaly detection model

Anomaly detection models are commonly used to identify events that have a low probability of occurrence but often cause fatal harm to the system (Chandola et al., 2009). Since crimes tend to be concentrated in specific areas, only a few street images are spatially associated with crime events (Weisburd, 2015). Therefore, the crime risk information contained within the street view images can be classified as anomally. Anomaly detection tasks that distinguish between anomalous and normal data is a One-Class Classification (OCC) tasks. Early OCC research focused on using statistical methods for feature extraction and developing classifiers. Since 2017, deep learning methods have become the mainstream of OCC research (Perera et al., 2021) which have made progress in several areas, such as cybersecurity intrusion detection (Kim & Kim, 2021), medical pathology image detection (Schlegl et al., 2017), One approach to deep learning-based OCC is to learn normal features and compare differences between test data and normal features, with greater variations indicating anomalous data (Ruff et al., 2018).

The convolutional neural network (CNN) structures can be utilized to achieve high accuracy in anomaly detection algorithms for images (Minhas & Zelek, 2019). Cohen and Hoshen (2020) utilized pyramidal neural networks to detect anomalous images and localize anomalous parts. Massoli et al. (2021) proposed the MOCCA framework to extract features at different depths of deep neural networks, thereby enhancing network discrimination in single-classification (OCC) problems. Sabokrou et al. (2018) introduced the first single-classification model based on GAN networks, which enhances the interpreter's normalization ability while iteratively reconstructing features.

3. Materials and method

Fig. 1 depicts a flowchart illustrating the pickpocketing crime risk assessment with coupled street view images using deep anomaly detection. The methodology comprises three fundamental stages: (1) Data preparation. Collected the 2018 judgment documents using a web crawler and subsequently extracting the crime locations using a natural language processing models, the crime locations were spatially with the street view images; (2) Mapping urban crime risk using the proposed Constructed Crime Anomaly Detection framework based on Street View (CADSV). We evaluated the risk of pickpocketing crimes by calculating image feature similarities between street view image; (3) Model interpretability analysis. Used POI kernel density data to characterize the drivers of the pickpocketing crime risk. This analysis utilized the Random Forest and SHAP models for interpretability. Additionally, we explored whether these drivers are consistent with the objective environmental risks characterized by the Street View imagery.

3.1. Study area and data

Shenzhen (Fig. 2) is a typical migrant city and the most developed city in South China, consisting of 10 districts. There are significant differences in economic development between downtown and suburban areas in Shenzhen (Meyer, 2016). The downtown areas are Shenzhen's political, economic, and cultural centre, including Nanshan District,

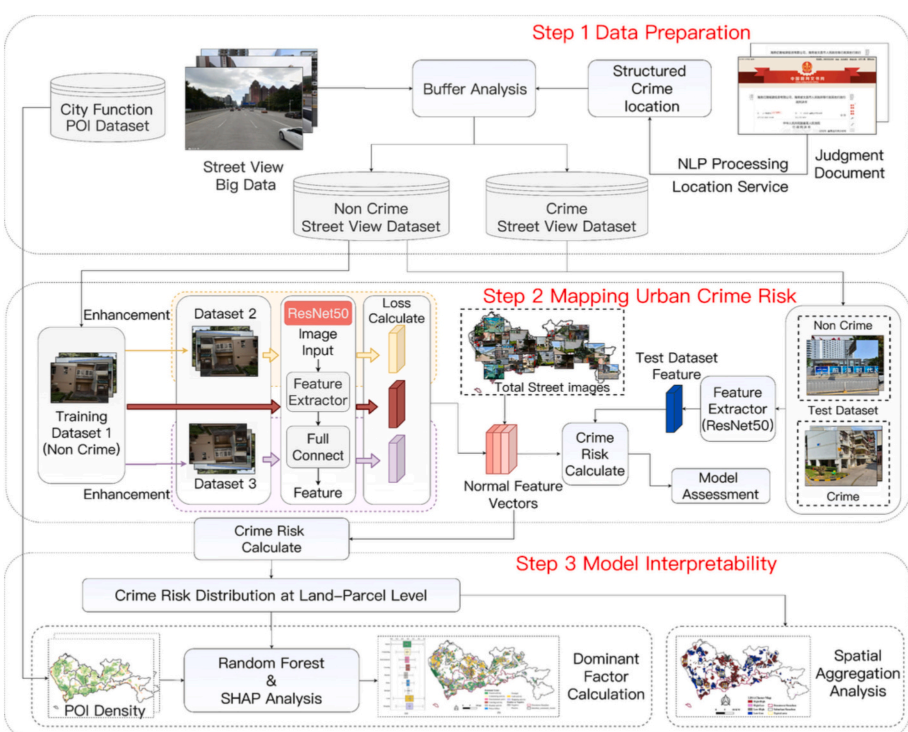

Fig. 1. Schematic overview of pickpocketing crime risk assessment with coupled street view images using deep anomaly detection.

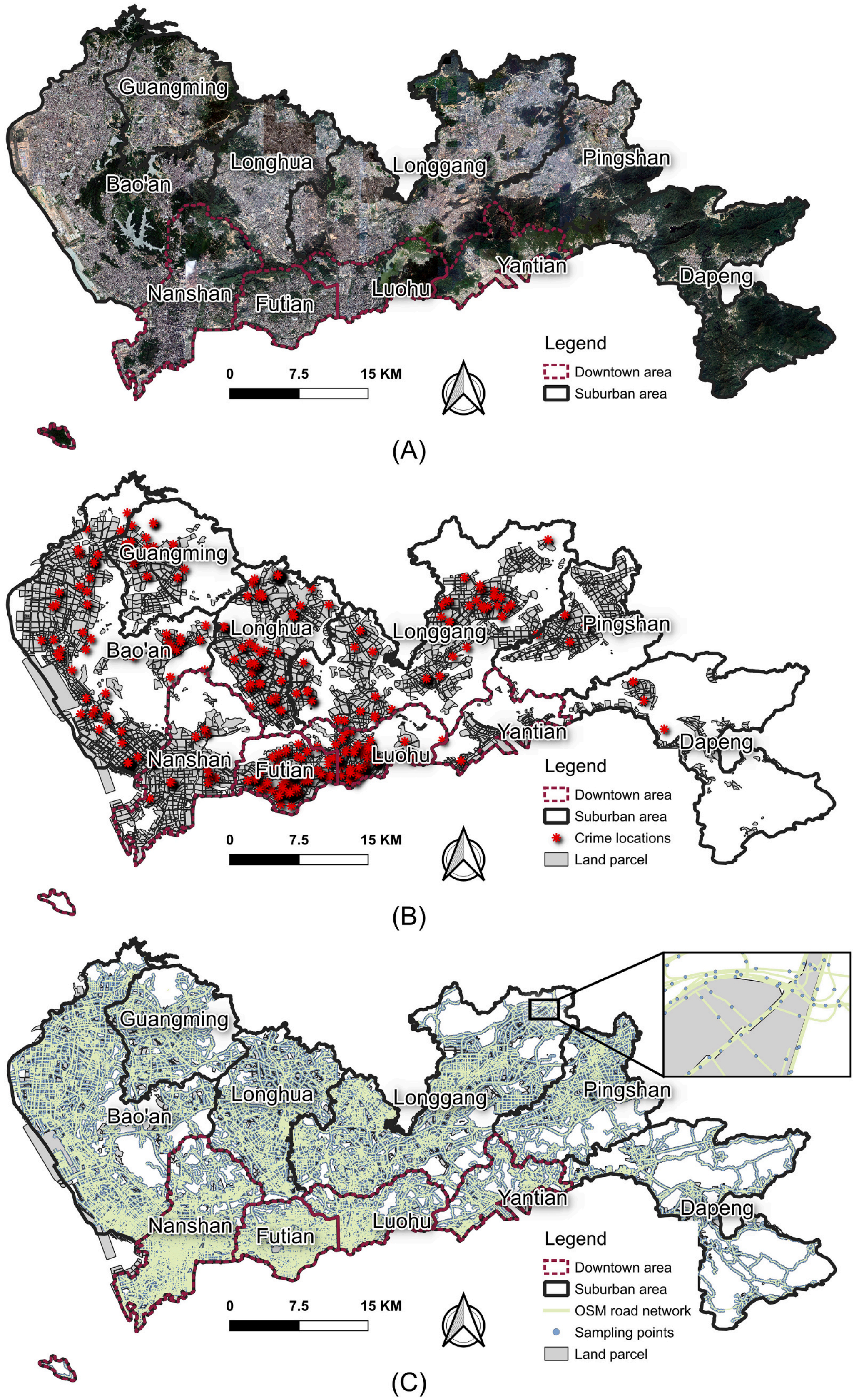

Fig. 2. (A) High-spatial resolution remote sensing imagery and (B) Crime locations and land-parcels in the study area (Shenzhen). (C) Road network and street view image sampling points.

Futian District, Luohu District, and Yantian District. Meanwhile, the suburban areas comprise Longhua District, Longgang District, Pingshan District, Dapeng District, Guangming District, and Baoan District, with a complex composition of foreign and migrant populations. It should be noted that the administrative division of Shenzhen underwent significant reorganization in 2018. To ensure study validity and offer support for future research efforts, all data used in this study were collected from 2018.

Land use planning parcels represent the fundamental unit of urban cadastral management in China. This study employed Shenzhen land use parcel data as the primary analytical unit. These parcels consist of 6913 records retrieved

from the Shenzhen Planning and Natural Resources Bureau website (https://pnr.sz.gov.cn/).

Acquiring crime locations from social media platforms has been explored in literature by Hipp et al. (2019). However, this approach may not provide credible crime data and therefore needs to be supported by police or official documents. In this study, the pickpocketing data were obtained from the China Judicial Documents website (http://wenshu.court.gov.cn). The Supreme People's Court of China mandates that all Chinese courts to publish judgment documents on the web, including information such as the cause, time, and location of the crime. To validate the accuracy of this data, previous studies have analyzed crime cases from different fields (Cai & Xin, 2019; Miao et al., 2016). Our study captured all criminal cases from 2018, which amounted to 7535 cases. Of these, pickpocketing accounted for 9.05 % (or 682) of all sentencing documents. Natural language processing models were utilized to extract pickpocketing crime locations, which are accurate up the building level or street level can be spatially matched with street view images.

Street View Image has been utilized in prior studies to reflect the physical environment or residents' perceptions of cities (Helbich et al., 2019; Wang et al., 2019b; Zhang et al., 2018). In this study, Baidu Street View images from 2018 were employed to depict the urban environment in Shenzhen. As one of the largest street view service providers available in China, Baidu Street View covers a vast majority of Chinese cities (Kang et al., 2020). This study used the road network data of Shenzhen city in 2018 were obtained via OpenStreetMap using the OpenStreetMap API. Byun and Kim (2022) have noted that acquiring street views at the street level with a distance of 200 m can effectively reveal the urban environment and its changes. There are also studies on crime that use 200 m as a buffer range for data collection, an interval that is considered to best describe the scale of urban communities (Kadar et al., 2016). Therefore, this study employed a sampling approach that involved the collection of road network data at 200 m intervals, thus obtaining a total of 38,717 sampling points. Subsequently, street view images were acquired from four horizontal directions (0°, 90°, 180°, and 270°) to simulate human visual perception. In total, 154,868 street view images were obtained for all sampling points. These images were then labeled as either pickpocketing risk images or non-pickpocketing risk images based on crime locations. Consistent with the sampling interval, a buffer radius of 200 m was selected to identify street images at risk of pickpocketing. Consequently, a total of 2712 street images were flagged as being at risk of pickpocketing. All other street view images were classified as normal images. It is worth noting that each street view image was labeled only once, although some street view images were located within buffer zones of multiple crime locations.

Point of Interest (POI) data have been demonstrated to effectively reflect the socioeconomic and functional structural characteristics of cities (Yao et al., 2017). In this study, POI data were utilized to analyze the relationship between pickpocketing crimes and urban functions at the micro-scale. The POI data used in this study were derived from Gaode Map (https://www.amap.com/), one of China's largest online map providers. A total of 213,476 POI data points from the year 2018 were collected in the study area, which were classified into five major categories: Catering & Entertainment, Education & Health care, Industry, Finance & Insurance, and Other Five major categories (Hu & Han, 2019). These categories were further divided into nine second-level categories, which are Life Services (39,590, 18.55 %), Transportation (37,536, 17.58 %), Landscape (2515, 1.18 %), Police (2544, 1.19 %), Medical Institutions (20,327, 9.52 %), Restaurants (75,030, 35.15 %), Finance (14,165, 6.64 %), Entertainment (17,107, 8.01 %) and Shopping Malls (4662, 2.18 %). We calculated the density of each type of POI density using kernel density analysis.

3.2. Extracting crime information by treating it as the anomalies

3.2.1. The propose of the assumption and the overall framework

Limited availability of crime data in certain regions makes it difficult to associate street view images with criminal activity. Typically, crimes tend to occur in specific locations according to the crime concentration theory (Weisburd, 2015). To address this issue, we propose that crime information can be viewed as anomalies within urban landscapes. Based on this assumption, we developed a Crime Anomaly Detection based on Street View (CADSV) framework for mining pickpocketing risk information from spatially sparse street view images and performing large-scale risk mapping. The framework is threefold (Fig. 3): 1) The normal feature vectors extraction. 20 % of street view images (totally 29,744 images) labeled with non-crime were randomly selected. The ResNet-50 Network was used to extract the normal feature vectors for all street view image. Normal feature vectors include the feature vector extracted for each image. 2) Verify the effectiveness of the extracted normal feature vectors for revolving crime information. 3) Mapping the crime risk for all street view images in the study area.

In this study, we aimed to evaluate the effectiveness of our proposed Crime Anomaly Detection based on Street View (CADSV) framework. To achieve this objective, we randomly selected 29,744 street view images to extract normal feature vectors using the ResNet-50 Network. Subsequently, we selected 10,148 street images to assess the performance of these extracted normal feature vectors.

It is important to note that we included all crime-labeled images in the test set, resulting in a total of 2712 such images. To ensure accurate assessment of the capability of the extracted features in assessing crime risk, we selected four times as many normal-labeled images as crime-labeled ones. Thus, we randomly selected 7436 images for this purpose. It should be noted that there were no strict guidelines for selecting this number; however, we considered 10,148 images to be sufficient for the evaluation process.

3.2.2. Normal feature extraction and feature adaptation

Self-supervised deep anomaly detection is considered a One-Class Classification (OCC) problem. However, when the amount of data is limited, a trained Convolutional Neural Network (CNN) may not effectively capture the semantic information within image dataset, resulting in suboptimal performance. Recent studies have demonstrated that pretraining can improve the effectiveness of model in deep anomaly detection. The CNN network is trained in a larger dataset to get the original feature vectors, which are then adapted features for use in the target dataset.

Feature adaptation aims to map the data from a different source and target domains into a feature space such that they are as similar as possible to each other in that space. Contrast learning is an excellent and effective self-supervised learning method commonly used for the feature adaptation of pre-trained feature extractors (Khosla et al., 2020; Reiss & Hoshen, 2023). The contrast learning method optimizes the prediction task by extracting a dataset $x^{'}$ of batch size N from the training set and training it against the data-enhanced $x^{'}$ with the loss function shown in Eq. 3.1. Where ϕ denotes the feature extraction module used to compute the feature vectors, which in this study represents the ResNet-50 model with normalization added at the last fully connected layer. This is because this method speeds up the convergence of the model and ensures adaptive normalization of the feature data (Yao et al., 2021a).

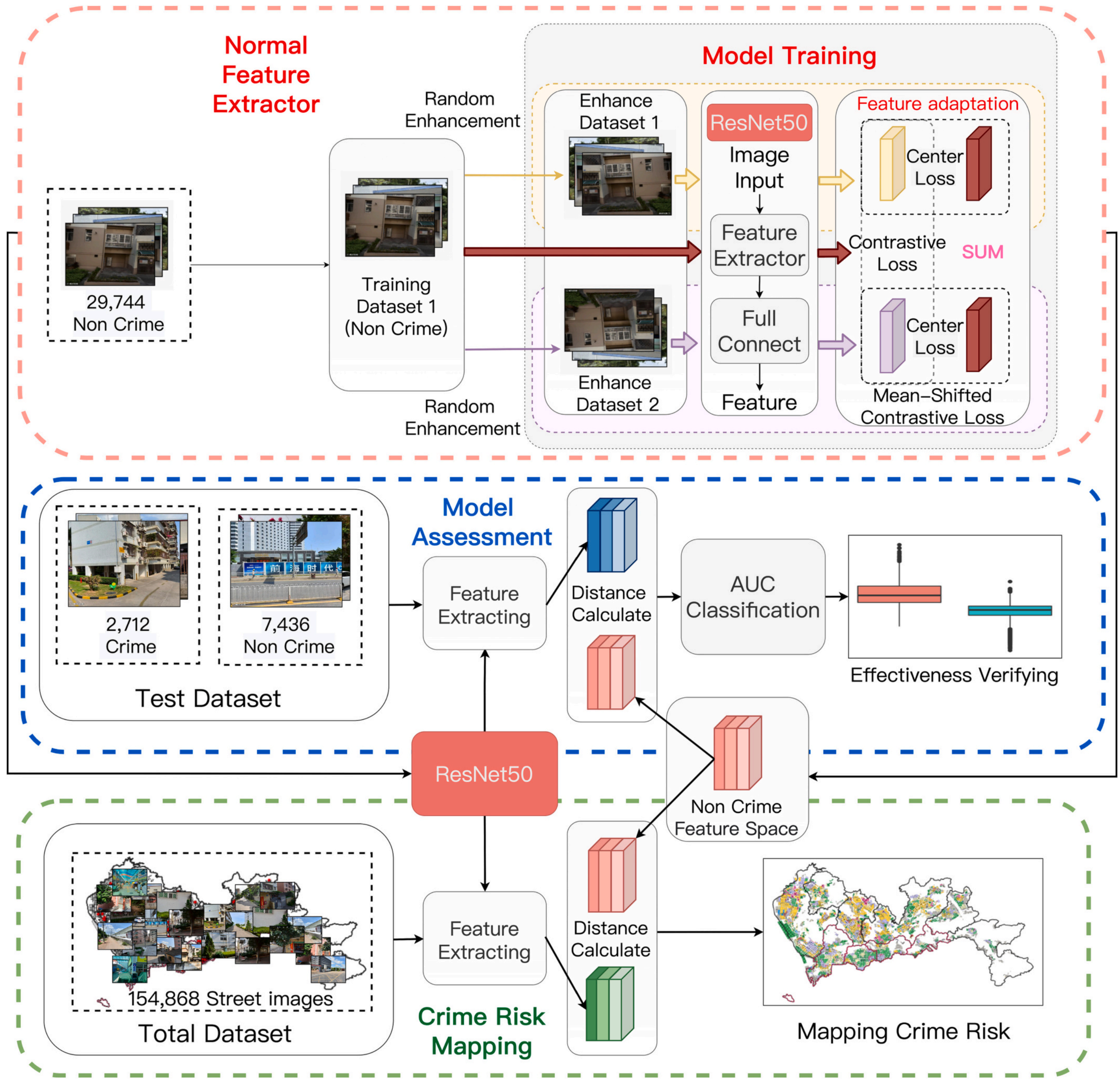

Fig. 3. Crime Anomaly Detection based on Street View (CADSV) framework coupled with Resnet-50 and MSAD.

$$L_{contrastive}(x_{i}^{'},x_{i}^{'}) = -log\frac{exp((\phi(x^{'})\cdot\phi(x^{'}))/\tau)}{\sum_{i=1}^{2N}\mathbb{1}[x_i \neq x^{'}]\cdot exp((\phi(x^{'})\cdot\phi(x_i))/\tau)} \tag{3.1}$$

The temperature hyperparameter, which is utilized in contrast learning to regulate the strength of penalty for negative samples (Wang & Liu, 2021), is denoted by τ in Eq. 3.1.

However, although the above contrast learning method is very effective for feature adaptation, for deep anomaly detection in OCC, it may lead to catastrophic collapse, where the accuracy of prediction decreases instead as the number of training increases. Therefore, we used a newly developed loss function, Mean-shifted contrastive loss, proposed by Reiss and Hoshen (2023). It is shown that this solves the problem of dimensional collapse that may occur in the field of image anomaly detection and surpasses the latest previous models in OCC classification. The objective function of mean-shifted loss is shown in Eq. 3.2, and c_{train} denotes the normalized centre of all training images:

$$\theta(x) = \frac{\phi(x) - c_{train}}{\|\phi(x) - c_{train}\|} \tag{3.2}$$

The method is not only able to calculate the Euclidean distance difference between the feature vector of a single image x and and c_{train}, but also normalizes the sample difference to the unit sphere and maximizes the distance between negative and positive samples. Besides, in order to reduce the distance between the $x^{'}$ samples and c_{train} after data enhancement, we also introduced the angular center loss (ACL), and the formula is shown in Eq. 3.3:

$$L_{angular}(x) = -\phi(x)\cdot c_{train} \tag{3.3}$$

To sum up, the objective function used in this study that combines the above two constraints is shown in Eq. 3.4:

$$L_{msc}(x',x'') = -log\frac{exp((\theta(x')\cdot\theta(x''))/\tau)}{\sum_{i=1}^{2N}\mathbb{1}[x_i \neq x']\cdot exp((\theta(x')\cdot\theta(x_i))/\tau)} + L_{angular}(x') + L_{angular}(x'') \tag{3.4}$$

In this study, small batches of data after data enhancement from the original training set are represented by x' and x''. The data enhancement method includes a series of ways such as flipping, cropping, and Gaussian filtering of the original image features, and ensures that the data enhancement results for x' and x'' are not the same by introducing randomness.

In this study, the ImageNet dataset was selected to pre-train the ResNet-50 network. And the feature adaptation was conducted using the proposed objective function in Eq. (3.4). Each image in the training set was extracted with a feature vector of 2048 dimensions. Normal feature vectors include the feature vectors extracted for each image.

3.2.3. Risk scoring and feature assessment

The process of anomaly scoring (risk scoring) for a street image from the test set is depicted in Fig. 4. Firstly, the pre-trained ResNet-50 network was utilized to extract the feature vector of the test image. Secondly, the KNN model was used to find the K nearest normal vectors for the test feature vector. The K was selected as two, and Euclidean Distance was used as distance metric, referring to the previous work (Reiss et al., 2021). Third, the risk score of the test image was calculated as the cosine distance between the test feature vector and the two nearest normal vectors. The cosine distance can take into account both the Euclidean distance and the angular distance between the features. It has a better deep anomaly detection performance than the other two distance metrics (Reiss et al., 2021). The scoring formula is shown in Eq. 3.5:

$$s(x) = \sum_{\phi(y)\in N_k(x)} 1 - \phi(x)\cdot\phi(y) \tag{3.5}$$

where $N_k(x)$ denotes the k features in the training set that have the closest cosine distance to $\phi(x)$. Moreover, the training set consists of street scenes where no crime occurred, while $s(x)$ indicates the distance of the input street scenes from the normal street scenes. This value ranges between 0 and 1, with higher values indicating a greater probability that the input streetscape belongs to the pickpocketing area.

The Anomaly Scoring was conducted on each image in the test set to assess the ability of the extracted normal feature vectors to characterize non-criminal features. The scoring results of normal-labeled street view images were compared with crime-labeled street view images to verify the effectiveness of the normal vectors.

The model's performance was evaluated using the AUC metric, which represents the area under the ROC curve (Ling et al., 2003). This metric can address classification result biases towards the majority class when the sample data is imbalanced (Burez & Van den Poel, 2009). The ROC curve is a probability curve that plots the true positive rate (TPR) against the false-positive rate (FPR), while the AUC measures the model's ability to classify correctly. An AUC close to 1 indicates good separability, while an AUC of 0.5 implies no category separation ability.

In this study, the best threshold for classifying whether street view images contain pickpocketing risk features or not was determined using the Youden index (Schisterman et al., 2005; Youden, 1950). Classification accuracy was subsequently assessed using Recall and F1-score. Recall can indicate the proportion of positive samples being correctly predicted in the classification results. On the other hand, the F1-score provides a comprehensive evaluation of classification model accuracy and recall. The formulas used to calculate Recall and F1-score are as follows:

$$Recall = \frac{TP}{TP + FN} \tag{3.6}$$

$$F1 = \frac{2*Precison*Recall}{Precision + Recall} \tag{3.7}$$

where TP denotes the number of correctly identified pickpocketing street view images and FN denotes the number of incorrectly predicted as normal street view images.

3.3. Interpretability analysis based on random forest and Shapley

This study employed the CADSV framework to calculate the street-level pickpocketing crime risk score for each image. The average risk score of street view images within each land parcel was used to characterize the crime risk score in that particular parcel. To investigate the effects of different POIs on the pickpocketing risk, this study utilized a SHAP method to interpretability analysis the results. SHAP (SHapley Additive Explanations) is a data feature analysis method based on game theory (Lundberg & Lee, 2017). This interpretable model that can integrate multiple variables effectively and reveal the contribution of each input spatial data in the model. The SHAP model finds wide application across various fields such as crime (Xie et al., 2022; Zhang et al., 2022) and medicine (Kim & Kim, 2022; Yao et al., 2022). The SHapley values (Štrumbelj & Kononenko, 2014) were calculated in the SHAP model to interpret the contribution and influence of the input

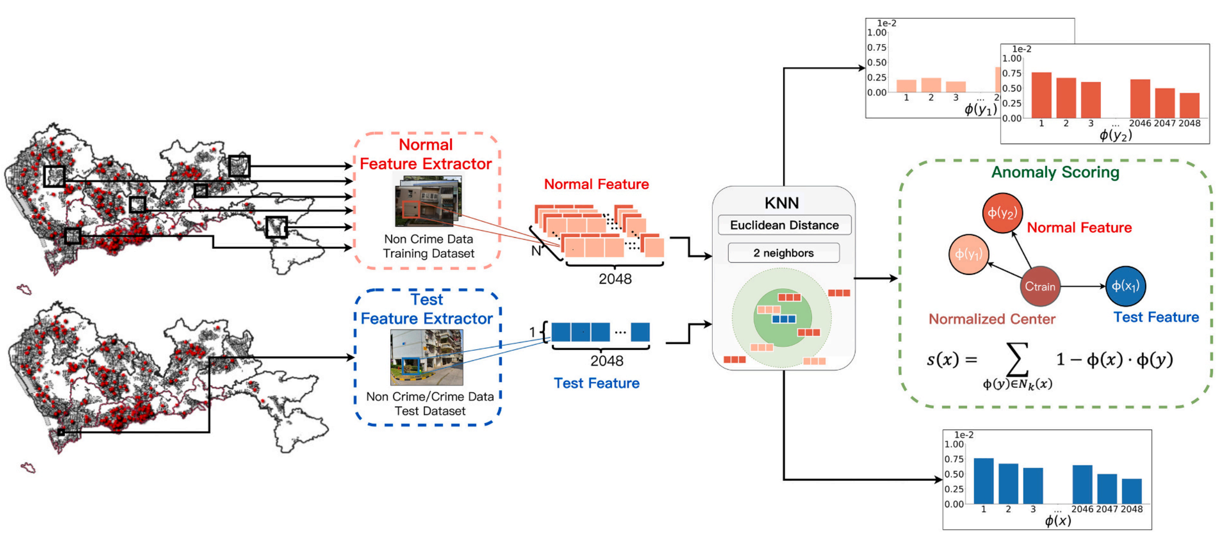

Fig. 4. The process of anomaly scoring for one test image.

features. Specifically, this study employed the Shapley model is used to explain the degree of contribution of different POIs to crime risk. The formula for calculating SHapley values is presented in Eq. 3.8.

$$\phi_i = \sum_{S\subseteq N\setminus\{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!}(v(S\cup\{i\})-v(S)) \tag{3.8}$$

where $\phi_i$denotes the SHapley value of the i-th independent variable and $|N|$ is the number of POI types; S represents an arbitrary disjointly arranged subset of each POI attribute except the i-th variable; $v(S\cup\{i\})$ denotes the output of the model when all data appear; $v(S)$denotes the output of only the input subset model. Following the above method, the SHapley value of each feature can be calculated by sequentially arranging and sampling each multi-source spatio-temporal data.

In addition to explaining the contribution of spatial variables using the SHAP method, this study used a random forest (RF) model to fit the POI density of different types within each land parcel to the pickpocketing indices. RF model has been used to analyze complex nonlinear correlations between variables in spatial analysis (Hengl et al., 2018; Nussbaum et al., 2018; Rodriguez-Galiano et al., 2012). It can effectively avoid correlation issues in high-dimensional features and has been shown to be the most effective nonlinear fitting model in previous studies (Fernández-Delgado et al., 2014).

4. Result

4.1. Model accuracy

The normal feature vectors were extracted from the training dataset and evaluated in the test dataset. A 5-fold cross-validation was conducted to obtain the hyperparameters of learning rate (0.0005), batch size (64), and epoch (150) were obtained for the training dataset. The risk scores of street view images in the test dataset are shown in Fig. 5. The test dataset include 7436 normal-labeled images and all 2712 crime-labled images. The results revealed a significant difference in risk scores between crime-labeled and normal-labeled images, with values of 0.41 and 0.29, respectively. Accuracy assessment shows that the AUC, Recall, and F1-Score were 0.921, 0.816, and 0.767, respectively. The scoring result in test dataset demonstrate that the extracted feature vector effectively characterizes the normal urban landscape (Fig. 5) and can detect crime information as anomalies.

Fig. 5c shows the percentage of criminal and non-criminal images in the test set at different intervals of crime risk values. When the crime risk values were less than approximately 0.36, the percentage of non-crime images greatly exceeded that of crime images in each interval of risk values. Conversely, when the crime risk values were >0.36, the proportion of crime-risk images rapidly increased and surpassed that of

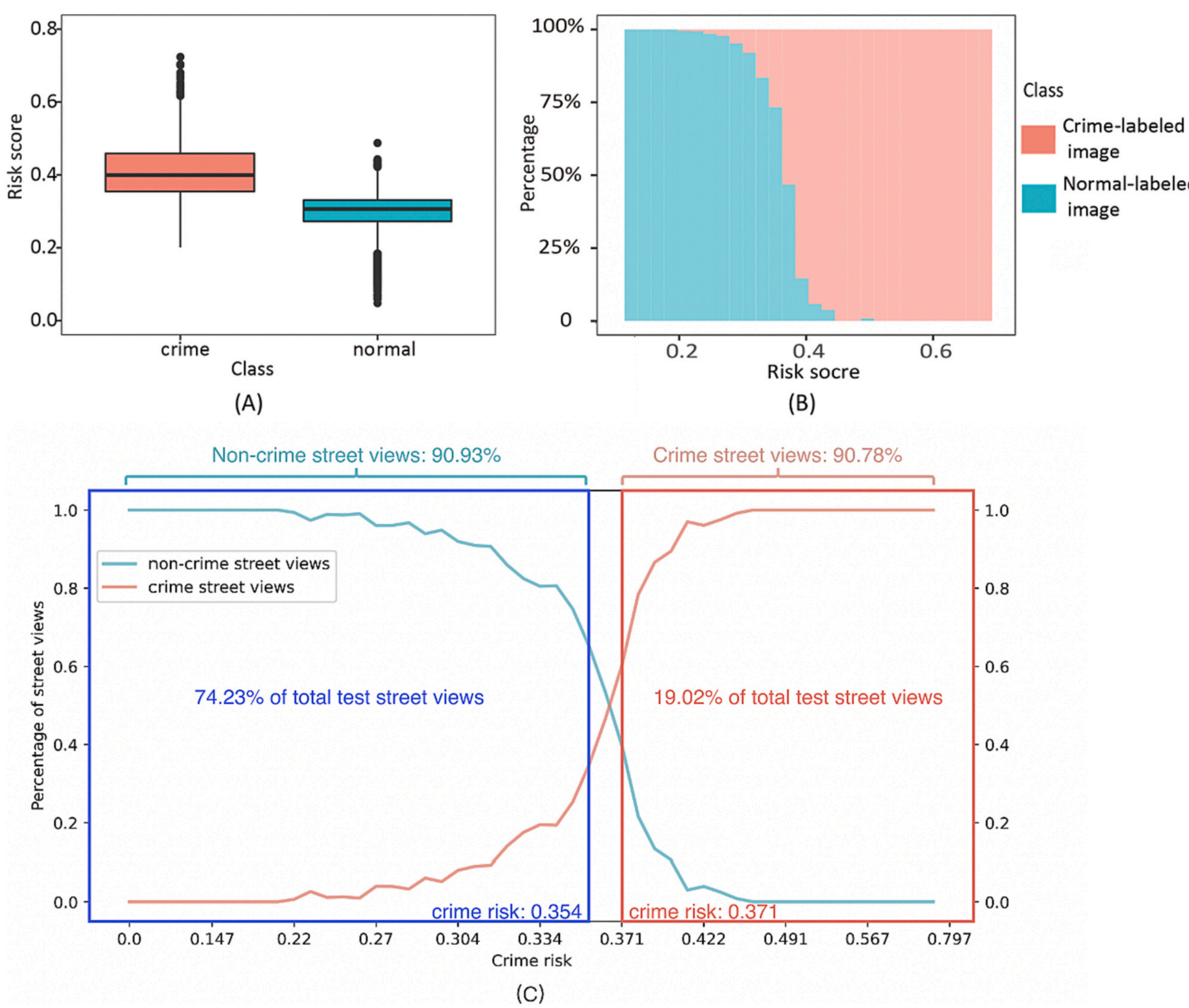

Fig. 5. Risk scores in test datasets: (A) box plot; (B) Percent stacked histogram; (C) Line chart of the percentage of crime/ non- crime street views in each crime risk value interval.

non-crime risk. Therefore, the crime risk values obtained from this model exhibit good discrimination between crime and non-crime images. Further statistical analysis revealed that when the crime risk was <0.354, including images with a ratio of 74.22 % of cases, 90.05 % of them were non-criminal. In contrast, when the crime risk was >0.371, 19.04 % of the images were included, while 90.78 % were criminal. At a crime risk value of 0.354–0.371, the ability to distinguish crime images from non-crime images was found to be the weakest, with a ratio of 276:409 between crime and non-crime images in this range. However, this represents only a small percentage (6.75 %) of the total number of images. In conclusion, the crime risk value can distinguish between crime and non-crime images well.

Having established the validity and rationality of the extracted normal feature vectors, pickpocketing crime risk scoring was conducted for all street view images. The resulting scores from street view images were then aggregated into land parcel level. To investigate the driving factors of pickpocketing crime risk, this study used the random forest to fit the nonlinear relationship between each type of POI feature and the pickpocketing crime risk index. In the fitting step, the random forest out-of-bag samples were randomly accounted for 30 %, and the number of decision trees (estimators) was set to 400. The R^2, RMSE, and MAE were 0.455, 0.016, and 0.868, respectively. These results demonstrated that socioeconomic features revealed by POI data could effectively explain the mapping results of the pickpocketing risk mapping result in most areas.

4.2. Parcel-scale pickpocketing risk mapping

This study mapped the risk distribution of pickpocketing crime for all land parcels in Shenzhen (Fig. 6). The results revealed a high pickpocketing risk in the central city with an average value of 0.362, and a low pickpocketing risk in the peripheral city with an average value of 0.347. This observation suggests that commercial and transportation activities, which are more prevalent in the central urban areas, may play a significant role in shaping the risk pattern of pickpocketing crime in Shenzhen.

Fig. 7 shows the street view images and their corresponding risk scores of typical functional zones. In general, for each functional area, the more dense the urban building bias, the more chaotic and disorganized the visual perception of the environment, the more likely pickpocketing is to occur. For instance, although the risk of crime inside a factory is low, while a construction site underway is at greater risk. Our findings indicate that the risk of pickpocketing crime is higher in areas relatively disorganized and underdeveloped areas. Tangwei Urban Village, a shantytown in Shenzhen, with a high migrant population and weak security management, had a higher risk of pickpocketing crime (0.404) compared to the resident community Yijing Community (0.354), as seen in Fig. 6. The Business Centre generally had good infrastructure, but its dense flow of people and disorder led to a higher risk of pickpocketing crime. For example, Mixc World Shopping Mall, one of the major business centres in Shenzhen, Mixc World Shopping Mall had a higher risk of pickpocketing crime (0.378) than the average value (0.351). These observations highlight the importance of considering visual perception when evaluating pickpocketing crime risks in urban settings.

The study results highlight variations in pickpocketing crime risk risks across functional areas. Tourist attractions showed a high average pickpocketing risk value (0.404) due to Shenzhen's well-developed tourism industry, with numerous scenic green spaces attracting many tourists and providing accessible targets for criminals (Fig. 7). Additionally, the risk values of pickpocketing in residential (0.372), commercial (0.362), industrial (0 0.368), and school (0.357) areas were higher than the average pickpocketing risk value (0.351). Such functional areas were characterized by dense crowds that offered an opportunity for disorder, making offenders more likely to commit pickpocketing crimes. Parks and landscape spaces were adjacent to residential areas also had many open spaces (Fig. 7 Scenery) that attracted offenders.

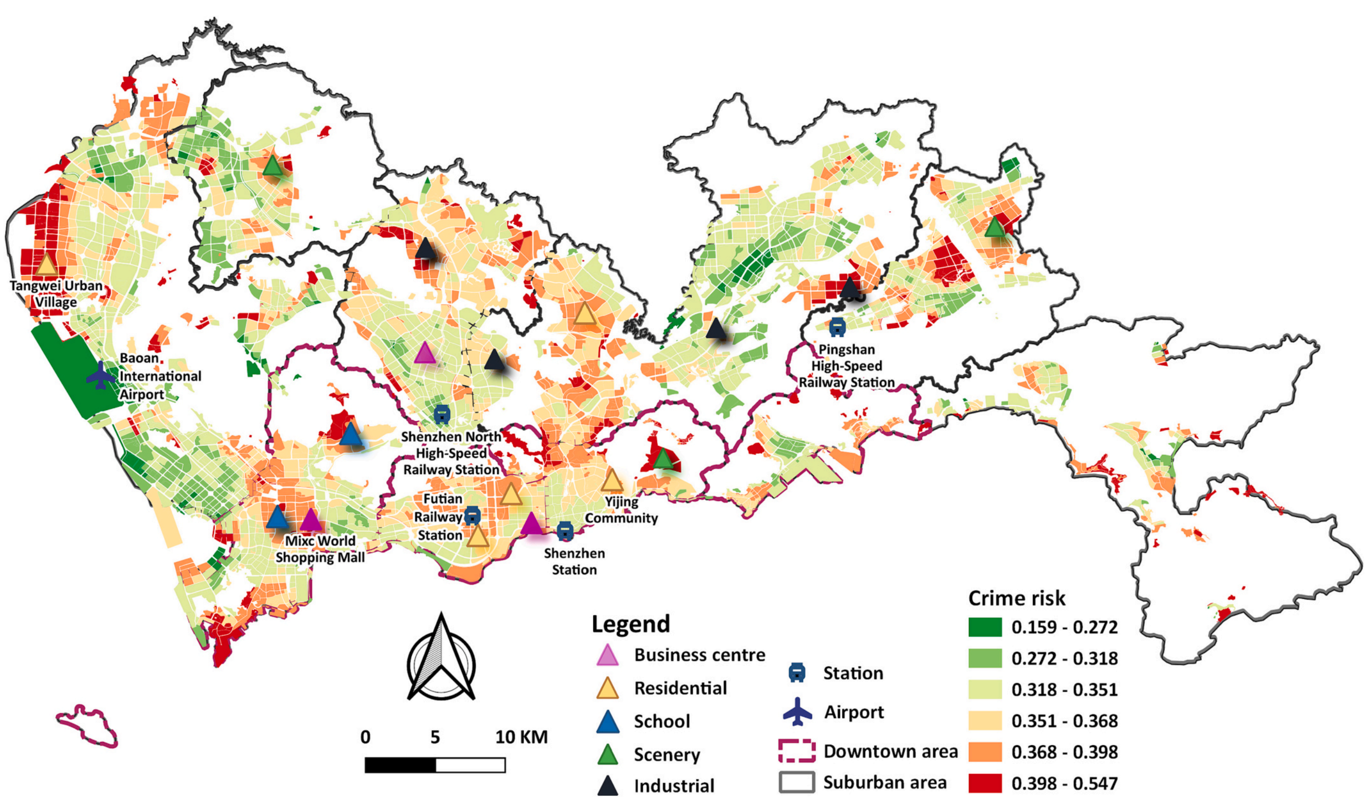

Fig. 6. The distribution of pickpocketing crime risk at land parcel-level in Shenzhen. The triangle marks typical functional areas.

Fig. 7. Street view images of typical functional zones of Shenzhen: The Crime risk axis represents the pickpocketing risk score assessed by the CADSV model, and the Land-use types axis represents typical functional zones.

Moreover, differences in social structure within functional areas can lead to heterogeneity in pickpocketing crime risk. As shown in Fig. 6, the transportation facilities in the city centre had a 39.4 % higher average pickpocketing risk than transportation facilities in other areas. Specifically, transportation facilities such as Shenzhen Station (0.377) and Futian Railway Station (0.373), with an average risk value of 0.375, while those in Baoan International Airport (0.172), Pingshan High-Speed Railway Station (0.331), and Shenzhen North High-Speed Railway Station (0.306), had an average pickpocketing risk level of 0.269. As shown in Fig. 7, the streetscape of stations was very similar in different areas. The risk was higher in the city centre with better economic development, suggesting that economic factors play plays a complex role in pickpocketing crime.

4.3. Spatial aggregation analysis of pickpocketing crime risk

The global Moran' I index of pickpocketing crime risk in Shenzhen was 0.591 (p-value <0.001, z-score 51.219), indicating a significant spatial correlation. Furthermore, local spatial autocorrelation analysis based on the Local Moran's I index was conducted to investigate the pattern of urban crime aggregation (Fig. 8). The results indicated that social factors significantly influenced the clustering pattern of pickpocketing crimes. Approximately 29.9 % of areas in Shenzhen had high-high aggregation of a pickpocketing crime risk, which were mainly located in urban central areas, such as University Town (Fig. 8(A)) and Futian CBD (Fig. 8(B)). These regions were characterized by a high concentration of people and wealth, making them prime targets for pickpocketing crimes. Additionally, areas outside the central urban area, such as Tangwei Urban Village (Fig. 8(C)), also showed high-high aggregation due to their inadequate infrastructure construction and a large

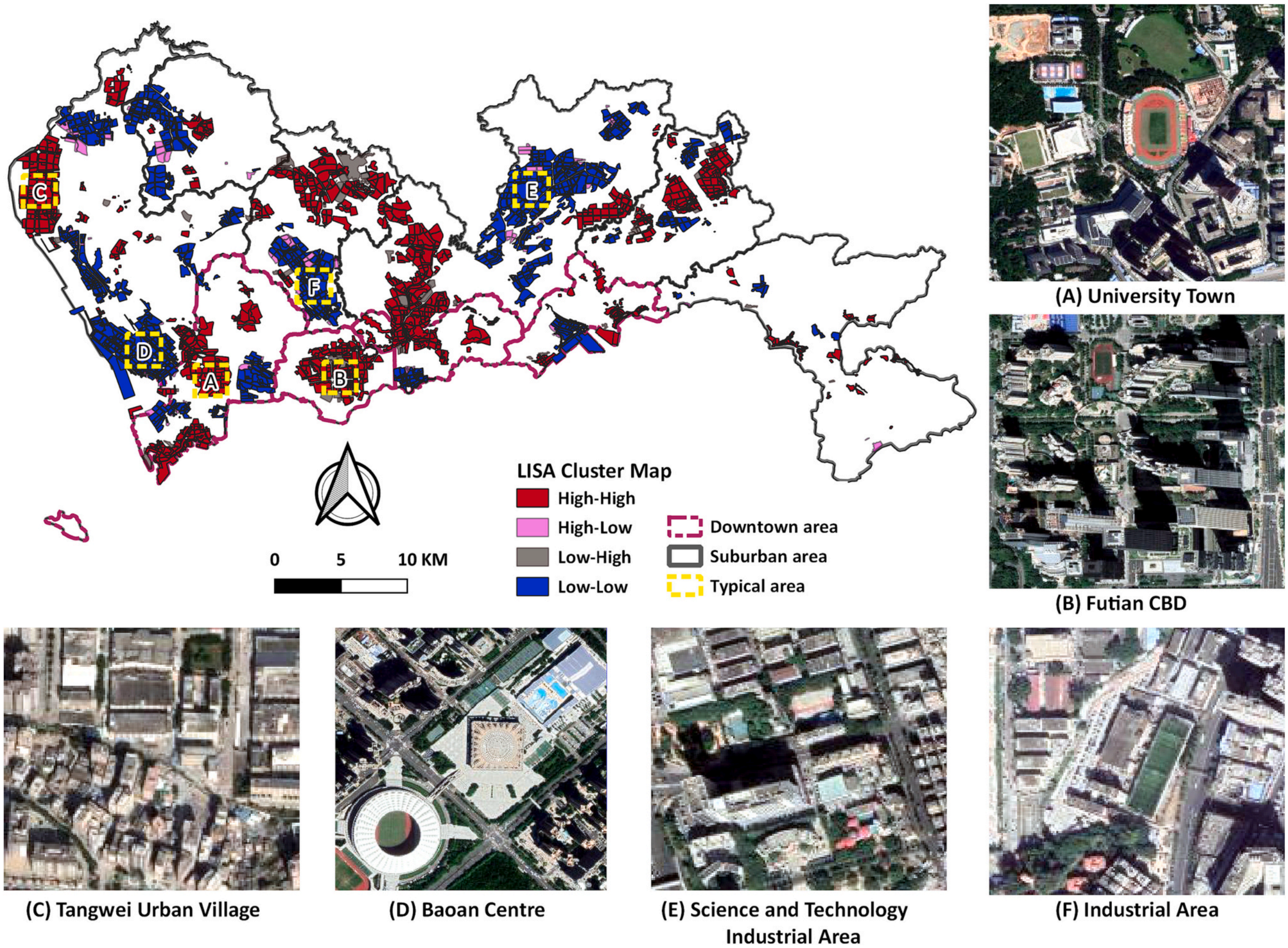

Fig. 8. Results of parcel-scale pickpocketing crime risk aggregation in Shenzhen and remote sensing images of some typical areas.

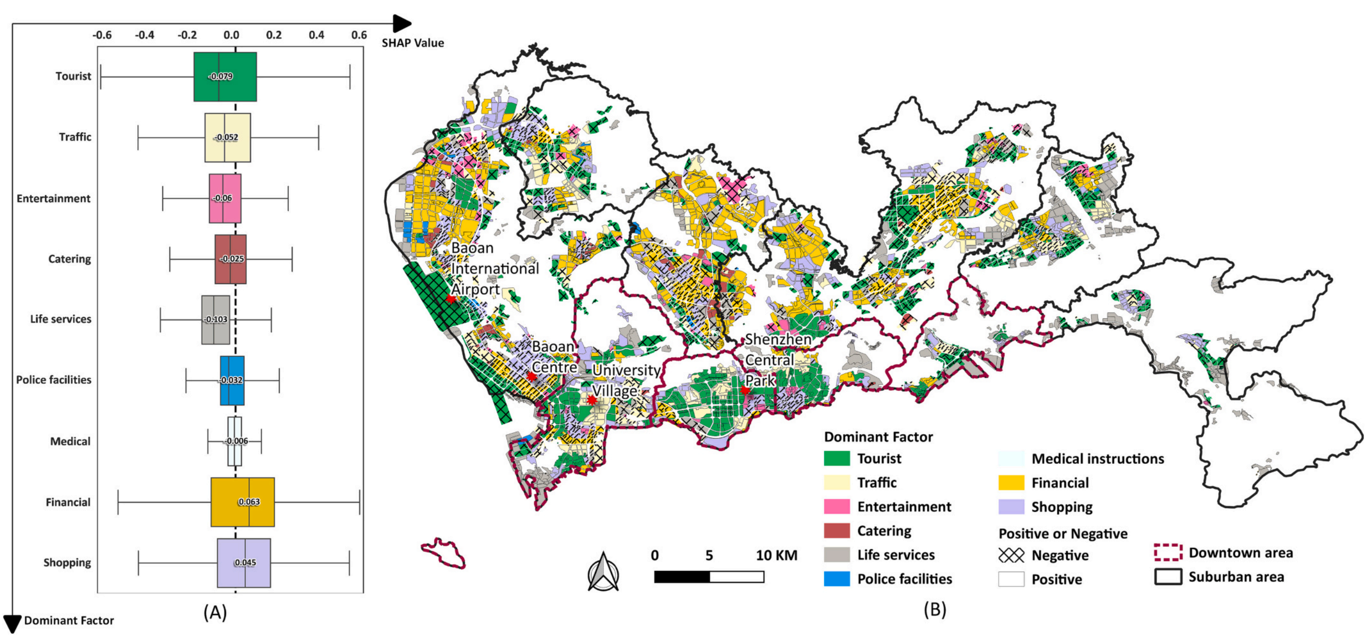

Fig. 9. Distribution of SHAP values and main drivers for all land parcels in Shenzhen: (A) shows the statistical analysis of SHAP values for all land parcels with the same characteristics; the SHAP Value axis indicates the magnitude of SHAP values, and the Dominant Factor axis indicates each type of POI characteristic that affects the risk of pickpocketing; (B) shows the drivers that have the greatest impact on the risk of pickpocketing in the parcel, which is obtained based on the average absolute magnitude of SHAP.

proportion of migrant residents.

Conversely, 26.8 % of the area exhibited low-low aggregation regions, mainly distributed outside the central city. Notably, these regions typically had better built environments and overall policing, evident in areas such as the central southern region of Baoan district, the science and technology industrial area in the south-central Longhua district, and the northeastern industrial area in Longgang district (Fig. 8(D)-(F)). Our study findings indicate socioeconomic conditions and the built environment in neighbouring regions greatly influence the spatial pattern of crime risk. The results are consistent with the hypothesis proposed by Sparks (2011a).

4.4. Explainable spatial distribution of pickpocketing risks

After fitting the relationship between each type of POI feature and the risk of pickpocketing crime, we calculated the SHAP values for each type of POI feature (Fig. 9(A)). Residents' routine activity was mainly carried out in five facilities, namely Traffic, Entertainment, Catering, Shopping, Financial, and Life services facilities (Boivin, 2018). The results indicate that routine activity was the most important factor influencing pickpocketing crime risk and positively correlates with the risk of pickpocketing crime. Compared to Tourist (−0.052), Medical Institutions (−0.003), and Police (−0.009), routine activity had the greatest impact on the risk of pickpocketing crime with an overall SHAP value of 0.079. According to the resident's routine activity theory (Cohen & Felson, 1979), Routine activity facilities provide criminals and potential targets that are prone to criminal activity. Tourist characteristics had a negative impact on pickpocketing crime risk scores, with a negative median SHAP value (−0.079). The Tourist feature reduced the risk score in many samples, which indicates that tourism had a high potential to reduce the risk of pickpocketing crime in urban areas. The results are consistent with the findings of (Bogar & Beyer, 2016) that urban landscape features are associated with reduced crime rates.

The present study investigated the determinants of pickpocketing crime risk on all parcels in Shenzhen Shenzhen, revealing insights into the spatial heterogeneity of crime risks (Fig. 9(B)). The influence of routine activities on the risk of pickpocketing crime was found to be spatially heterogeneous. In the downtown area, routine activity facilities were observed to positively affect the risk of pickpocketing (average SHAP value: 0.184). For example, in the vicinity of Shenzhen University Town, where traffic was the main driver, urban residents moved around for work and education purposes, and the population was more mobile, increasing the risk of pickpocketing crime in the area. However, residents' routine activity facilities in suburban areas negatively impact crime risk (average SHAP value: −0.118). For example, Shopping and Financial features dominated the dominant factors near the central area of Baoan district, which reduced the crime risk of the area. This may associate with increased guardianship (Boivin, 2018), which stabilizes social order. These results support the hypothesis that routine activity may increase or decrease criminal activity (Boivin, 2018).

The present study revealed the heterogeneous effect of tourist attractions on the risk of pickpocketing crime. Tourist (average SHAP value: 0.098) had a predominantly positive effect on crime risk in economically developed urban areas. For instance, Shenzhen Central Park witnessed an increased risk of pickpocketing due to the congregation of many tourists, making tourism the main driver of crime risk in the area (Zhong et al., 2011). In contrast, tourists in suburban areas (average SHAP value: −0.119) mainly negatively affect crime risk. For example, crime risk in Baoan International Airport was driven by tourism, but it reduced the risk of pickpocketing crime in the area. This could be attributed to the presence of a large area of public green spaces near the airport, which stabilizes social order and reduces the risk of pickpocketing crime (Jennings & Bamkole, 2019). Our study shows the uncertainty of Tourist's effect on crime risk in the region due to socioeconomic influences, which is in line with the findings of (Groff & McCord, 2012).

5. Discussion

The extraction of crime information from street view images reflecting the built environment is essential for urban governance and crime risk analysis. However, the number of street view images labeled as crime occurred is often very less. This issue is particularly true in China since the most reliable crime data source is the Chinese judgment documents, which do not contain all criminal cases. This study is an active attempt to extract crime risk through urban built environments using spatially sparse crime data. To achieve this, we adopted an alternative approach by evaluating the distribution of pickpocketing crimes based on OCC-based anomaly detection and street view images. In addition, it is the first exploration of the relationship between the built environment and pickpocketing crime risk in a large Chinese city. Previous studies have analyzed cities' physical environment and socioeconomic characteristics as reflected in street view images in several U. S. cities and explored their relationship with urban crime rates (He et al., 2017; Zhang et al., 2021).

5.1. Interpretation of the findings

In this study, we developed a Crime Anomaly Detection based on the Street View (CADSV) model, which can effectively extract deep semantic information from massive street view for assessing pickpocketing risks. Our results show that the street view images can accurately reflect the city's physical environment and provide reliable assessments of the risk of pickpocketing crimes, as confirmed by the high accuracy of the CADSV model (AUC = 0.921, recall = 0.816). Through comparative and explainable analysis, we obtained a micro-scale pickpocketing risk distribution map of the study area and confirmed the reliability of the results.

Our findings reveal that pickpocketing crime in Chinese megacities exhibits strong spatial autocorrelation, consistent with previous studies' observations that crime tends to be concentrated in small areas (Groff et al., 2010; Weisburd et al., 2004). The observed decrease in crime risk decreases with distance from the downtown provides quantitative support for the social disorganization theory proposed by Shaw et al. (1942). In the downtown area, pickpocketing is a high prevalence and aggregation of pickpocketing crime are significant due to dense human traffic, making it challenging to manage. Simultaneously, in the course of ongoing urban expansion due to population and economic growth, the influx of migrant workers and the gradual deterioration of the physical environment in older urban areas (Li et al., 2014) have become the dominant drivers of the high prevalence of pickpocketing crimes in urban village areas (Liu, 2010). In contrast, the suburbs, and industrial parks outside the central city, where a large number of immigrants live and work (Roitman & Phelps, 2011), display a better built environment and exhibit low-low aggregation of pickpocketing crimes. These findings highlight the complexity of the impact of urban function on pickpocketing crime. They also confirm that confirm that the complex roles of the urban physical environment, neighborhood socioeconomics, and migrant population all play a significant role in shaping the spatial distribution of pickpocketing crime risk in Chinese cities (Sparks, 2011b).

This study also utilized the Random Forest and SHAP model to interpretively analyze the relationship between pickpocketing crime, urban function, and urban environment at the microscopic scale. The proposed model achieved high accuracy ($R^2 = 0.455$) and reliability (RMSE = 0.016) by employing urban functions to fit pickpocketing risk, thus quantitatively confirming the crucial role of different urban functions in shaping regional pickpocketing risk. Consistent with the routine activity theory proposed by Cohen and Felson (1979) and the crime pattern theory proposed by Brantingham and Brantingham (2013), our findings highlight that routine activity in the central city is a critical factor that enhances pickpocketing risks. Furthermore, we observed that the high intensity of economic activity in commercial areas contributes significantly to crime incidence in China.

The results demonstrate that both the built environment information and urban functional information captured in street view images play a role in shaping crime incidence. Regarding the built environment, areas with high-density urban buildings and visually chaotic and disorganized surroundings are more susceptible to pickpocketing crimes, while areas with better-organized environments have lower risks. Concerning urban functional zones, densely populated areas that are challenging to fully secure and where high-value items are prevalent pose a higher risk of pickpocketing crimes. Examples include high-traffic attractions, shopping centres, isolated factories or residential areas, schools with a high number of minors, among others. We can use these findings to guide urban planning and security management. For example, in high-traffic areas such as commercial centres and tourist attractions, surveillance and security forces can be strengthened in advance to reduce the threat of crime by installing additional warning signs and alarm facilities based on unusual risk situations reflected in the street view images. In important areas such as residential and school zones, residents and students can be encouraged to exercise extra vigilance towards high-value property. Additionally, open spaces like parks and attractions could benefit from increased police patrols and surveillance equipment deployments in crowded areas can be increased to improve security perceptions and prevent pickpocketing and other crimes from happening. Furthermore, functional areas may also reflect differences in social structures. For instance, the transport facilities present a higher risk of crime in the city centre than in other areas, which may have complex links to social phenomena such as frequent movements of movement of people, conflicts arising from intricate social structures, and the allocation of police forces. In the long term, governments should to promote social development and long-term security by improving the social structure and raising the level of the economy.

The above discussion underscores the multifaceted nature of pickpocketing crime in China's megacity, which cannot be explained by a single theory of crime. The occurrence of such crime is influenced by a combination of regional economic development, urban physical environment, and routine activity, among other factors. Moreover, interpretable analyses have uncovered complex spatial heterogeneity in the drivers of pickpocketing crime across China's megacity. Routine activity exerts a positive impact on areas with intense human activity areas in downtown regions but a negative impact in suburban areas. Similarly, tourist activity positively affects crime risk in urban centers but has a negative effect in the suburbs. These findings indicate that the dichotomy of China's urban-rural structure, characterized by the differences in physical space, industrial infrastructure, and economic composition across regions (Ann et al., 2015; Long et al., 2016), gives rise to significant variations in the underlying drivers of crime.

As a developing country, China faces the challenge of operating with a relatively constrained police force and fewer police services available per capita compared to developed countries, resulting in inadequate law enforcement resources to combat pickpocketing crimes (Hyland & Davis, 2019; Wang et al., 2014). To enhance policing effectiveness, our findings suggest deploying police patrols strategically high-risk areas for pickpocketing crimes while implementing video surveillance systems in urban villages, shopping centres, and economic activity centers. Furthermore, increasing anti-pickpocketing campaigns at daily activity locations such as bus stops and metro stations could raise residents' security awareness and contribute to reducing pickpocketing risks. Our study further highlights that urban villages with significant migrant populations are at greater risk of pickpocketing crime. Therefore, improving service facilities in urban villages and providing more employment opportunities may aid in enhancing urban policing efforts.

There are difficulties in mapping the real spatial pattern of crime risk due to the specificity of data collection for the judgment document. The anomaly detection model in this study learns from sparse data about hidden crime risks and finds a spatial mismatch between the number of crime events and the risk of the area. Current policy-making authorities quantify the level of policing in an area mostly based on government survey data, such as judgment documents. Conclusions based on such data may therefore lead to problems such as misallocation of public resources and misguided business investments. Our findings may provide support to government policy makers or commercial investors.

5.2. The spatial mismatch between the judgment document and mapped crime risk

Our study has revealed a spatial mismatch between crime risk and the original crime data obtained from judgment document, as depicted in Fig. 10(A). This disparity is a tangible manifestation of the sparse and biased sampling problem that this research has explored. Specifically, the spatial distribution of data collected through sentencing instruments is exceedingly sparse and closely associated with factors such as population density and law enforcement efficiency, rendering it difficult to accurately reflect the actual spatial pattern of crime risk.

In practice, acquiring a judgment document involves a lengthy process comprising three stages: (1) commission of a crime and successful theft; (2) notification of the police by the victim or public body, leading to the opening of a case and arrest of the suspect; and (3) filing of a case and commencement of prosecution against the accused by the victim or public body in a court of law. Consequently, the data we collect for each judgment paper represents not only the occurrence of a crime, but also the diligence of the court and the efficiency of the police in executing the case. Regarding the distribution of crimes, since not all criminal incidents go through the aforementioned process, the sentencing paper data can only serve as a sparse sample point and cannot directly depict the full scope of criminal activities.

Regarding the sparsity of sampled data, Table 1 presents the area covered by crime points and the number of street view images under varying buffer distances. The results reveal that with a buffer distance of 200 m, only 3 % of the region is labeled as crime-related, with an average of 4.17 street view images per buffer. If a shorter buffer distance of 50 m is chosen, then merely 0.2 % of the area is covered, and each buffer can only include 0.31 street image. Given the limited proportion of relevant data, it becomes arduous to offer a comprehensive and accurate spatial pattern at a global level.

With regards to the biased nature of the data, we counted the judgment instruments for all cases (including cases in which the location is not publicly available) in each administrative region of Shenzhen in 2018 based on the data provided by the Judgment Instruments website (https://wenshu.court.gov.cn/), as shown in Fig. 10(B). The figure depicts darker colors indicating a higher total number of judgment instruments within each respective region. It is evident that the aggregation level of crime points obtained through our opportunity judgment instruments closely aligns with the number of judgment instruments generated by courts in each region.

In order to further validate our previous assertion concerning human activity, we have gathered Real-time Tencent user density (RTUD) data. Tencent is one of the largest internet companies in China, with a user base exceeding 800 million individuals utilizing its diverse range of internet services. Through Tencent Maps or WeChat, when users engage in location-related activities, their relevant location information is recorded, enabling RTUD data to capture population distribution during specific periods. The raw data is stored as a raster image format comprising of 24 bands that represent each hour of the day. This study utilizes an overlay of the 24-h average change in population density over weekdays to generate a graph (He et al., 2020). A correlation can be observed between higher crime spots and elevated levels of people's activities, once again affirming the notion that data acquisition does not accurately reflect the complete volume of criminal incidents.

This study has examined areas of mismatch to establish a connection between Fig. 10 (A) and Fig. 8. Specifically, in Figs. 10 (A-a) and Fig. 8 (C), it is apparent that despite the limited sample of pickpocketing crime events in Tong Mei Urban Village, the village exhibits a high risk of

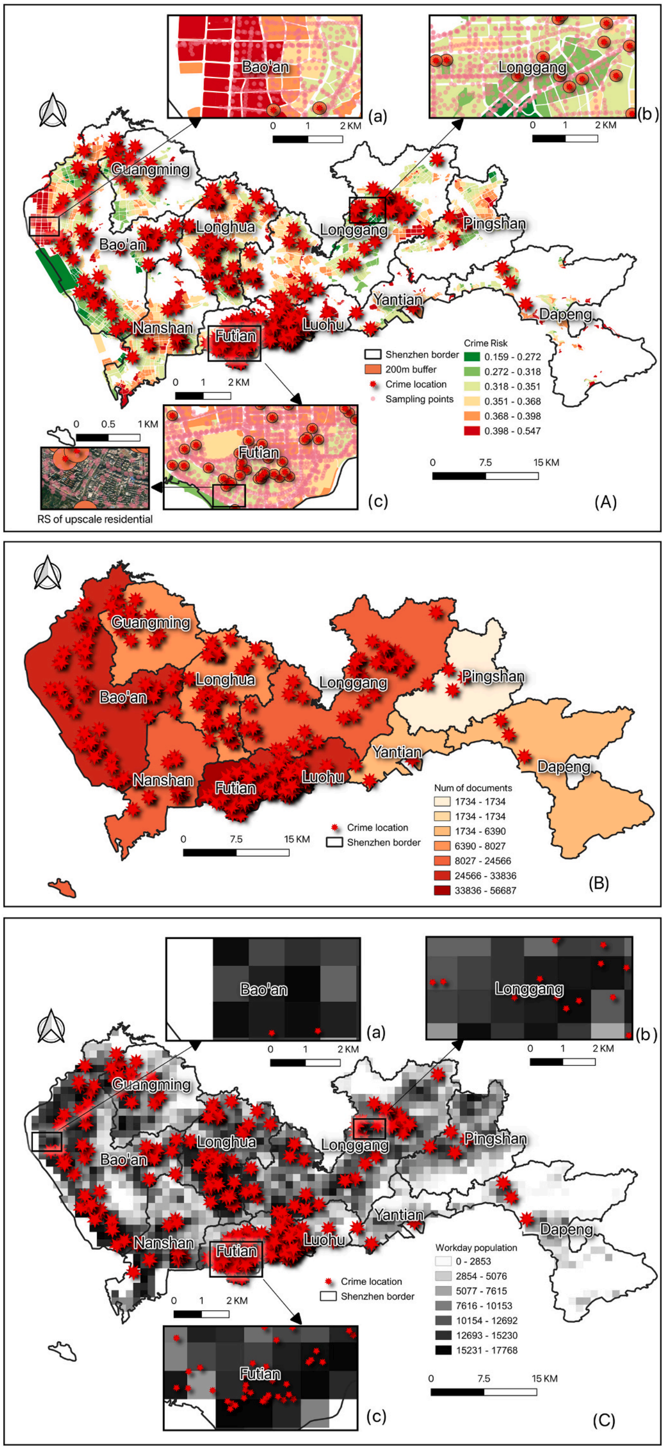

Fig. 10. The overlap of pickpocketing events between the judgment document and: (A) the crime risk distribution; (B) the total number of judgment documents of each district; (C) The average population distribution of Shenzhen during the working day.

criminal activity due to its inadequate infrastructure and chaotic building structure. Conversely, Fig. 10 (A-b) and Fig. 8 (E) demonstrate that emerging technology industrial parks and factories have a lower risk of crime, attributed to their well-organized and tidy environment. Further, Fig. 10 (A-c) corresponds to Fig. 8 (B), where the Futian Central Business District displays a medium-high risk of criminal activity due to the concentration of wealth. However, there are also areas with a low risk of criminal activity, as seen in Fig. 10(A-c). These regions are primarily located in residential areas within the city center, which may be due to enhanced security facilities and the higher quality of residents.

Table 1
The area and images that are covered by crime event according to different buffer.

Buffer range	50 m	100 m	150 m	200 m	250 m	300 m
Area percentage covered by buffer zones	0.2 %	0.8 %	1.9 %	3.0 %	5.4 %	7.8 %
Average number of images covered by one buffer	0.31	1.36	2.51	4.17	6.69	9.45

5.3. Limitation and future works

The present study does have some limitations that must be acknowledged. First, the objective of this study is to utilize sparse crime data to reflect implicit crime risks in urban built environments using the technique of anomaly detection. However, it is important to note that this study conducted a crime risk analysis rather than an estimation of actual crime rates. There are interactions and complex causal relationships between risk and crime rates. Understanding these relationships is critical to developing effective crime prevention and governance strategies, and requires the integration of multiple social, economic, cultural and individual dimensions. Future studies may collaborate with law enforcement agencies to analyze the relationship between street view images and real-time alarm data utilizing the framework proposed in this paper.

The second limitation of this study pertains to the crime data used in model design. The premise assumption of this study is that crime information can be viewed as an outlier in the urban landscape. Therefore, we first removed street view images spatially associated with a crime based on judgment documents. After that, we randomly selected a certain number of street view images for training purposes so as to obtain a normal feature vector. However, we cannot guarantee that no crime has occurred in those areas since judgment instruments may not always contain all relevant data. Our hypothesis was that by using deep learning for feature extraction using numerous images, we could eliminate the influence of crime information could be eliminated as much as possible. The precision validation results also show such effectiveness. To further improve the accuracy, subsequent studies should take into account the more prior knowledge and eliminate as much as possible the street view images where crime may be present to obtain the most effective feature vector. Moreover, exploring the impact of different street view image acquisition intervals on the results would be valuable. More frequent street view sampling has the potential to yield better results , and in the field of crime, how to choose the most suitable analysis interval is also a topic worth exploring (Ramos et al., 2021). This study has successfully constructed a framework illustrating the feasibility of anomaly detection for exploring crime risk. Subsequent research endeavors can build upon this framework to delve deeper into this topic.

The third limitation is we utilized POI data to interpret the result of crime risk mapping. Prior research has found that POI data can effectively reflect the characteristics of socioeconomic structure (Yao et al., 2017). Furthermore, we have carried out some work to demonstrate the strong relationship between street view images and several social and environmental factors such as urban economic level and urban population structure (Wang et al., 2021; Yao et al., 2021b). However, POI data and street view images can only be proxy variables of socioeconomic characteristics and urban environment. To explain the risk mapping result more accurately, future studies will introduce more detailed census, travel survey, and trajectory data. Additionally, econometric models may be used to analyze the temporal and spatial correlations between the multiple urban structures and criminal behaviours at the micro-scale. Finally, follow-up research can also expand the research scale by obtaining global street view datasets and analyzing the similarities and differences of criminal behavior drivers in different cities worldwide.

The solution proposed in this study carries substantial practical value, as street view images are easy to obtain, models can be easily migrated to other areas, and it can be utilized by lay users to quickly comprehend crime risk levels in a particular area of the city. For instance, we can develop a mobile application that enables users to swiftly assess safety status of a city neighborhood. Visitors arriving in an unfamiliar city can rapidly determine whether a specific alley is safe. While police crime statistics are typically the most trustworthy in such cases, official data may not always cover the entire area. In such scenarios, our approach can assist users in identifying and mitigating potential safety concerns.

6. Conclusion

This study aims to propose a solution for extracting crime risk information from the built environment using the limited crime-labeled street view images. We also try to prove the association between the human perception of the built environment and urban pickpocketing crimes in China. To achieve these objectives, we propose a pickpocketing risk assessment model that combines deep anomaly detection techniques to reveal the crime risk from street view images. The SHAP was introduced to conduct an interpretable analysis of urban functions and crime risks. Through spatial distribution analysis of pickpocketing crimes based on judicial documents, our proposed CADSV model accurately and reliably maps out a micro-scale pickpocketing risk distribution in Shenzhen. Our results indicate street view images can effectively assess pickpocketing crime risk in Chinese cities, and the crime risk has a strong spatial autocorrelation. Moreover, we demonstrate that pickpocketing crime in China is driven by complex factors such as regional economic development, physical urban environment, and daily activities. These findings provide valuable insights for policing deployment and city management strategies. Nonetheless, this study does not discuss the association between crime risk and actual crime rates. Moreover, the inclusion of finer-scale geographic big data could be considered to identify crime-related street view images in anomalydetection, thereby offering more prior knowledge about crime risk. Furthermore, a more comprehensive interpretation of the crime risk mapping results is necessary to analyze the correlation between various urban structures and criminal behavior at finer spatial and temporal scales.

CRediT authorship contribution statement

Yao Yao: Conceptualization, Methodology, Writing- Reviewing and Editing. **Anning Dong.**: Data curation, Writing- Original draft preparation. **Zhiqian Liu**: Visualization, Investigation. ***Ying Jiang:*** Supervision.: **Zijin Guo**: Software, Validation.: **Junyi Cheng:** Writing- Reviewing and Editing. **Qingfeng Guan:** Writing- Reviewing and Editing **Peng Luo:** Conceptualization, Methodology, Writing- Original draft preparation.

Fundings

This work was supported by the National Key Research and Development Program of China (Grant No. 2019YFB2102903), the National Natural Science Foundation of China (Grand No. 41801306 and 42171466), Alibaba Innovative Research Project (No. 20228670), and China Scholarship Council.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Ann, T. W., et al. (2015). The key causes of urban-rural conflict in China. *Habitat International, 49*, 65–73.
Ashihara, Y., & Riggs, L. E. (1983). *The aesthetic townscape*. Cambridge: MIT Press.
Bernasco, W., Block, R., & Ruiter, S. (2013). Go where the money is: Modeling street Robbers' location choices. *Journal of Economic Geography, 13*(1), 119–143.
Bernasco, W., Ruiter, S., & Block, R. (2017). Do street robbery location choices vary over time of day or day of week? A test in Chicago. *Journal of Research in Crime and Delinquency, 54*(2), 244–275.
Bogar, S., & Beyer, K. M. (2016). Green space, violence, and crime: A systematic review. *Trauma, Violence & Abuse, 17*(2), 160–171.
Boivin, R. (2018). Routine activity, population (S) and crime: Spatial heterogeneity and conflicting propositions about the neighborhood crime-population link. *Applied Geography, 95*, 79–87.
Bouma, H., et al., 2014. Automatic detection of suspicious behavior of pickpockets with track-based features in a shopping mall. In Optics and Photonics for Counterterrorism, Crime Fighting, and Defence X; and Optical Materials and Biomaterials in Security and Defence Systems Technology XI (International Society for Optics and Photonics), 92530F.
Brantingham, P., & Brantingham, P. (2013). Crime pattern theory. In *Environmental Criminology and Crime Analysis (Willan)* (pp. 100–116).
Brunton-Smith, I., et al. (2023). *Estimating the reliability of crime data in geographic areas* (CrimRxiv).
Buil-Gil, D., Moretti, A., & Langton, S. H. (2022). The accuracy of crime statistics: Assessing the impact of police data Bias on geographic crime analysis. *Journal of Experimental Criminology, 18*(3), 515–541.
Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. *Expert Systems with Applications, 36*(3), 4626–4636.
Byun, G., & Kim, Y. (2022). A street-view-based method to detect urban growth and decline: A case study of midtown in Detroit, Michigan, USA. *PLoS One, 17*(2), Article e263775.
Cai, T., & Xin, Y. (2019). Child trafficking in China: Evidence from sentencing documents. *International Journal of Population Studies, 4*(2), 1–11.
Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys, 41*(3), 15.
Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 588–608.
Cohen, N., & Hoshen, Y. (2020). Sub-image anomaly detection with deep pyramid correspondences. ***arXiv e-prints***. arXiv:2005.02357.
Deshotels, T. (2013). The declining criminal arts-the pickpocket. *The International Journal of Crime, Criminal Justice and Law/Serials Publication, 8*, 69–76.
Ding, N., & Zhai, Y. (2021). Crime prevention of bus pickpocketing in Beijing, China: Does air quality affect crime? *Security Journal, 34*(2), 262–277.
Du, F., et al. (2020). Predictive mapping with small field sample data using semi-supervised machine learning. *Transactions in GIS, 24*(2), 315–331.
Fan, N., et al. (2020). Digital soil mapping over large areas with invalid environmental covariate data. *ISPRS International Journal of Geo-Information, 9*(2), 102.
Fan, Z., et al. (2021). Evaluation of urban crime risk based on agent simulation model. In *2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE) IEEE* (pp. 648–651).
Farrell, G., & Bouloukos, A. C. (2001). International overview: a cross-national comparison of rates of repeat victimization. *Crime Prevention Studies, 12*, 5–26.
Fernández-Delgado, M., et al. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research, 15*(1), 3133–3181.
Giménez-Santana, A., Caplan, J. M., & Drawve, G. (2018). Risk terrain modeling and socio-economic stratification: Identifying risky places for violent crime victimization in Bogotá, Colombia. *European Journal on Criminal Policy and Research, 24*(4), 417.
Groff, E., & McCord, E. S. (2012). The role of neighborhood parks as crime generators. *Security Journal, 25*(1), 1–24.
Groff, E. R., Weisburd, D., & Yang, S. (2010). Is it important to examine crime trends at a local "Micro" level?: A longitudinal analysis of street to street variability in crime trajectories. *Journal of Quantitative Criminology, 26*(1), 7–32.
Hajela, G., Chawla, M., & Rasool, A. (2021). Crime hotspot prediction based on dynamic spatial analysis. *ETRI Journal, 43*(6), 1058–1080.
He, J., et al. (2020). Accurate estimation of the proportion of mixed land use at the street-block level by integrating high spatial resolution images and geospatial big data. *IEEE Transactions on Geoscience and Remote Sensing, 59*(8), 6357–6370.
He, L., Páez, A., & Liu, D. (2017). Built environment and violent crime: An environmental audit approach using Google street view. *Computers, Environment and Urban Systems, 66*, 83–95.
Helbich, M., et al. (2019). Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China. *Environment International, 126*, 107–117.

Hengl, T., et al. (2018). Random Forest as a generic framework for predictive modeling of spatial and Spatio-temporal variables. *PeerJ, 6*, Article e5518.

Hipp, J. R., et al. (2019). Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly, 36*(4), 718–748.

Hipp, J. R., et al. (2021). Measuring the built environment with Google street view and machine learning: Consequences for crime on street segments. *Journal of Quantitative Criminology*, 1–29.

Hossain, S., et al. (2020). Crime prediction using spatio-temporal data. In *International Conference on Computing Science, Communication and Security (Springer)* (pp. 277–289).

Hu, Y., & Han, Y. (2019). Identification of urban functional areas based on POI data: A case study of the Guangzhou economic and technological development zone. *Sustainability, 11*(5), 1385.

Hu, Y., et al. (2018). A Spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation. *Applied Geography, 99*, 89–97.

Hyland, S. S., & Davis, E. (2019). Local police departments, 2016: Personnel. In *Washington, DC: Bureau of Justice Statistics (BJS) US Dept of justice, Office of Justice Programs*. Bureau of Justice Statistics.

Jennings, V., & Bamkole, O. (2019). The relationship between social cohesion and urban green space: An avenue for health promotion. *International Journal of Environmental Research and Public Health, 16*(3), 452.

Jing, F., et al. (2021). Assessing the impact of street-view greenery on fear of neighborhood crime in Guangzhou, China. *International Journal of Environmental Research and Public Health, 18*(1), 311.

Kadar, C., Iria, J. E., & Cvijikj, I. P. (2016). Exploring foursquare-derived features for crime prediction in new York City. *KDD-Urban Computing WS, 16*, 10–1145.

Kang, Y., et al. (2020). A review of urban physical environment sensing using street view imagery in public health studies. *Annals of GIS, 26*(3), 261–275.

Khosla, P., et al. (2020). Supervised contrastive learning. *Advances in Neural Information Processing Systems, 33*, 18661–18673.

Kim, Y., & Kim, H. K. (2021). Cluster-based deep one-class classification model for anomaly detection. *Journal of Internet Technology, 22*(4), 903–911.

Kim, Y., & Kim, Y. (2022). Explainable heat-related mortality with random Forest and SHapley additive exPlanations (SHAP) models. *Sustainable Cities and Society, 79*, 103677.

Lafree, G., & Birkbeck, C. (2010). The neglected situation: A cross-national study of the situational characteristics of crime. *Criminology, 29*(1), 73–98.

Li, L. H., et al. (2014). Redevelopment of Urban Village in China–a step towards an effective urban policy? A case study of Liede Village in Guangzhou. *Habitat International, 43*, 299–308.

Li, Y., et al. (2016). Supplemental sampling for digital soil mapping based on prediction uncertainty from both the feature domain and the spatial domain. *Geoderma, 284*, 73–84.

Ling, C. X., Huang, J., & Zhang, H. (2003). AUC: A statistically consistent and more discriminating measure than accuracy. In *Proceedings of the 18th International Joint Conference on Artificial intelligence (Acapulco, Mexico: Morgan Kaufmann Publishers Inc.)* (pp. 519–524).

Liu, Z. G. (2010). The mechanism of crimes in villages-in-city: A case study of "T Village" in Shenzhen. *Criminal Research, 17*(6), 64–71.

Long, H., et al. (2016). The allocation and Management of Critical Resources in rural China under restructuring: Problems and prospects. *Journal of Rural Studies, 47*, 392–412.

Lundberg, S., & Lee, S. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems(Curran Associates Inc.)* (pp. 4768–4777).

Massoli, F. V., Falchi, F., Kantarci, A., Akti, Ş., Ekenel, H. K., & Amato, G. (2021). MOCCA: Multilayer one-class classification for anomaly detection. *IEEE Transactions on Neural Networks and Learning Systems, 33*(6), 2313–2323.

Meyer, D. R. (2016). Shenzhen in China's financial center networks. *Growth and Change, 47*(4), 572–595.

Miao, L., et al. (2016). Add to favorite get latest update discussion on the improvement of the crime of medical accident in China: Based on the analysis of the cases of China judgments online and Jianxue Li case. *Medicine and Philosophy : A, 37*(12), 3.

Minhas, M. S., & Zelek, J. (2019). Anomaly detection in images. *arXiv e-prints*, 1905–13147.

Nussbaum, M., et al. (2018). Evaluation of digital soil mapping approaches with large sets of environmental covariates. *Soil, 4*(1), 1–22.

Oswald, M., et al. (2018). Algorithmic risk assessment policing models: Lessons from the Durham HART model and 'experimental' proportionality. *Information & Communications Technology Law, 27*(2), 223–250.

Perera, P., Oza, P., & Patel, V. M. (2021). One-class classification: A survey. *arXiv e-prints*. arXiv:2101.03064.

Ramos, R. G., et al. (2021). Too fine to be good? Issues of granularity, uniformity and error in spatial crime analysis. *Journal of Quantitative Criminology, 37*, 419–443.

Reiss, T., et al. (2021). PANDA: Adapting pretrained features for anomaly detection and segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2806–2814).

Reiss, T., & Hoshen, Y. (2023, June). Mean-shifted contrastive loss for anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence, 37*(2), 2155–2162.

Rodriguez-Galiano, V. F., et al. (2012). An assessment of the effectiveness of a random Forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 67*, 93–104.

Roitman, S., & Phelps, N. (2011). Do gates negate the City? Gated Communities' contribution to the urbanisation of suburbia in Pilar, Argentina. *Urban Studies, 48*(16), 3487–3509.

Ruff, L., et al. (2018). Deep one-class classification. In *Proceedings of the 35th International Conference on Machine Learning (Proceedings of Machine Learning Research: PMLR)* (pp. 4393–4402).

Rumi, S. K., Luong, P., & Salim, F. D. (2019). Crime rate prediction with region risk and movement patterns. *CoRR*. abs/1908.02570.

Sabokrou, M., Khalooei, M., Fathy, M., & Adeli, E. (2018). Adversarially learned one-class classifier for novelty detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3379–3388).

Schisterman, E. F., et al. (2005). Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. *Epidemiology, 16*(1), 73–81.

Schlegl, T., et al. (2017). Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging (Cham: Springer International Publishing)* (pp. 146–157).

Shaw, C. R., et al. (1942). Juvenile delinquency and urban areas: A study of rates of delinquents in relation to differential characteristics of local communities in American cities. *University of Chicago Press, 49*(1), 100–101.

Sparks, C. S. (2011a). Violent crime in San Antonio, Texas: An application of spatial epidemiological methods. *Spatial and Spatio-temporal Epidemiology, 2*(4), 301–309.

Sparks, C. S. (2011b). Violent crime in San Antonio, Texas: An application of spatial epidemiological methods. *Spatial and spatio-temporal epidemiology, 2*(4), 301–309.

Steffensmeier, D., Zhong, H., & Lu, Y. (2017). Age and its relation to crime in Taiwan and the United States: Invariant, or does cultural context matter? *Criminology, 55*(2), 377–404.

Štrumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems, 41*(3), 647–665.

ToppiReddy, H. K. R., Saini, B., & Mahajan, G. (2018). Crime Prediction & Monitoring Framework Based on Spatial Analysis. *Procedia Computer Science, 132*, 696–705.

Wang, F., & Liu, H. (2021). Understanding the behaviour of contrastive loss. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2495–2504).

Wang, R., et al. (2019a). Using street view data and machine learning to assess how perception of neighborhood safety influences urban Residents' mental health. *Health & Place, 59*, 102186.

Wang, R., et al. (2019b). The linkage between the perception of Neighbourhood and physical activity in Guangzhou, China: Using street view imagery with deep learning techniques. *International Journal of Health Geographics, 18*(1), 18.

Wang, R., et al. (2021). The distribution of greenspace quantity and quality and their association with Neighbourhood socioeconomic conditions in Guangzhou, China: A new approach using deep learning method and street view images. *Sustainable Cities and Society, 66*, 102664.

Wang, Y., et al. (2014). Stress, burnout, and job satisfaction: Case of police force in China. *Public Personnel Management, 43*(3), 325–339.

Webb, V. J., et al. (2011). A comparative study of youth gangs in China and the United States: Definition, offending, and victimization. *International Criminal Justice Review, 21*(3), 225–242.

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology, 53*(2), 133–157.

Weisburd, D., et al. (2004). Trajectories of crime at places: A longitudinal study of street segments in the City of Seattle. *Criminology, 42*(2), 283–322.

Wilson, J., & Kelling, G. L. (1982). Broken windows: The police and neighbourhood safety. *The Atlantic Monthly, 249*, 29–38.

Xiao, J., & Zhou, X. (2020). Crime exposure along my way home: Estimating crime risk along personal trajectory by visual analytics. *Geographical Analysis, 52*(1), 49–68.

Xie, H., Liu, L., & Yue, H. (2022). Modeling the effect of streetscape environment on crime using street view images and interpretable machine-learning technique. *International Journal of Environmental Research and Public Health, 19*(21), 13833.

Yao, Y., et al. (2017). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science, 31*(4), 825–848.

Yao, Y., et al. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science, 33*(12), 2363–2384.

Yao, Y., et al. (2021a). Delineating urban job-housing patterns at a parcel scale with street view imagery. *International Journal of Geographical Information Science, 35*(10), 1927–1950.

Yao, Y., et al. (2021b). Delineating urban job-housing patterns at a parcel scale with street view imagery. *International Journal of Geographical Information Science*, 1–24.

Yao, Y., et al. (2022). Assessing myocardial infarction severity from the urban environment perspective in Wuhan, China. *Journal of Environmental Management, 317*, 115438.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35.

Yue, H., et al. (2022). Detecting people on the street and the streetscape physical environment from Baidu street view images and their effects on community-level street crime in a Chinese City. *ISPRS International Journal of Geo-Information, 11*, 151.

Zhang, F., et al. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning, 180*, 148–160.

Zhang, F., et al. (2020). Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems, 81*, 101478.

Zhang, F., et al. (2021). "perception Bias": Deciphering a mismatch between urban crime and perception of safety. *Landscape and Urban Planning, 207*, 104003.

Zhang, G. (2022). Mitigating spatial bias in volunteered geographic information for spatial modeling and prediction. In *New Thinking in GIScience* (pp. 179–190). Springer.

Zhang, G., & Zhu, A. (2018). The representativeness and spatial bias of volunteered geographic information: A review. *Annals of GIS, 24*(3), 151–162.

Zhang, G., & Zhu, A. (2019a). A representativeness heuristic for mitigating spatial bias in existing soil samples for digital soil mapping. *Geoderma, 351*, 130–143.
Zhang, G., & Zhu, A. (2019b). A representativeness-directed approach to mitigate spatial bias in VGI for the predictive mapping of geographic phenomena. *International Journal of Geographical Information Science, 33*(9), 1873–1893.
Zhang, X., et al. (2022). Interpretable machine learning models for crime prediction. *Computers, Environment and Urban Systems, 94*, 101789.
Zhao, X., & Tang, J. (2017). Modeling temporal-spatial correlations for crime prediction. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 497–506). Singapore: Association for Computing Machinery.
Zhong, H., et al. (2011). Spatial analysis for crime pattern of metropolis in transition using police records and GIS: A case study of Shanghai, China. *International Journal of Digital Contents Technology and Its Applications, 5*(2), 93–105.
Zhu, A. X., et al. (2018). Spatial prediction based on third law of geography. *Annals of GIS, 24*(4), 225–240.