

密级：公开

编号：17110654

# 中山大學

## 博士学位论文

居民安全感知的微环境影响要素、

作用机制及空间模拟

---

Microenvironmental Correlates, Influencing  
Mechanism, and Spatial Simulation of Residents'  
Safety Perception

---

学位申请人： 敬峰瑞

导师姓名及职称： 柳林教授

专业名称： 地图学与地理信息系统

2021 年 05 月 26 日

# 中山大学博士学位论文

## 居民安全感知的微环境影响要素、 作用机制及空间模拟

Microenvironmental Correlates, Influencing  
Mechanism, and Spatial Simulation of Residents'  
Safety Perception

专业名称：地图学与地理信息系统

学生姓名：敬峰瑞

指导老师：柳林教授

答辩委员会（签名）

主席：李军

委员：石萬同和陳楚輝 李海

二零二一年五月 广州



# **Ph.D. Dissertation**

**Microenvironmental Correlates,  
Influencing Mechanism, and Spatial  
Simulation of Residents' Safety Perception**

**Ph.D. Candidate: Jing Fengrui  
Supervised By: Professor Liu Lin**

**Sun Yat-sen University  
May 2021, Guangzhou**



## 论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本人论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期：2021 年 5 月 26 日

## 学位论文使用授权声明

本人完全了解中山大学有关保留、使用学位论文的规定，即：学校有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版；有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆、院系资料室被查阅；有权将学位论文的内容编入有关数据库进行检索；可以采用复印、缩印或者其他方法保存学位论文；可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

保密论文保密期满后，适用本声明。

学位论文作者签名：

导师签名：



日期：2021 年 5 月 26 日

日期：2021 年 5 月 26 日



## 摘要

居民安全感知，又称为治安安全感或被害恐惧感，与居民日常生活息息相关，影响着个体的身心健康和日常行为。我国政府近年来加大了社会治理力度，着重提升人民的幸福感、获得感、安全感。安全感受到个体因素和环境因素共同作用。因此，探讨环境与居民安全感知的关系，实现社区居民安全感的空间模拟，不仅有利于丰富和拓展安全感知的理论研究，也有利于指导具体的社会治理实践。已有的环境与安全感知关系研究存在一定的不足，主要包括以下方面：① 传统的安全感知相关环境因素分析缺乏对客观微环境要素的关注；② 客观环境和安全感知的关系机制分析中缺乏对间接效应的关注；③ 传统的安全感知及其环境影响因素的获取缺乏一套客观性强、实时性高、成本低的测量手段。

因此，本研究以广州市七个中心城区（越秀区、荔湾区、海珠区、黄埔区、天河区、白云区、番禺区）为研究对象，使用百度街景数据、犯罪数据、问卷调查数据、人口普查数据、POI 设施点数据、手机数据、遥感影像数据等，对环境与居民安全感知的关系和机制进行研究，并进行社区居民安全感知的空间模拟。本研究的主要研究成果如下：

(1) 街景图像微环境要素提取。基于百度街景地图，提出了“数据获取-要素提取-空间分析”的街景图像微环境要素提取算法流程。本研究基于编程语言采集了广州市研究区内的街景图像，利用机器学习技术进行了三项街景图像要素挖掘任务。依托广州超算云计算平台，完成了所有街景图像的要素提取，并基于遥感影像完成初步验证。研究得出，① 基于机器学习技术，可以建立基于环境失序理论的街景微环境失序指标体系。② 空间分布特征上，富裕度、美丽度、绿视率等要素分布在主城区及绿化水平较高的区域。主城区的建筑物比例值较高，天空开阔程度水平低。外围城区的建筑物比例值低，天空开阔程度水平高。③ 空间集聚特征上，富裕度、美丽度、绿视率等要素相关性较高，高值区域主要分布在主城区、二沙岛、公园及大学城区域。建筑物比例和天空开阔程度的集聚特征明显，且呈负相关状态。

(2) 街景微环境相关要素对居民安全感知的直接作用分析。以个体安全感知为因变量，一系列个体属性指标、社区结构特征、社区环境特征、社区微环境

特征为自变量，本研究使用 OLS 多元线性回归模型进行了一系列模型分析。研究得出，① 相较于街景美丽度、富裕度、垃圾度，街景绿视率是理想的社区环境失序程度的指标，是显著影响居民安全感知的微环境指标。② 加入一系列微环境要素的模型效果比没有加入微环境要素的模型效果要好，说明街景微环境要素的加入有利于完善居民安全感知的模型。③ 性别、感知的建成环境失序和社会环境失序、社会融入、社区商业设施比例、社区政府机构比例等一系列指标是显著影响居民安全感知的因素。

(3) 街景微环境对居民安全感知的间接效应分析。以个体居民安全感知为因变量，以特定街景微环境要素为自变量，以直接受害经历和非正式社会融入为调节变量，以建成环境失序为中介变量，其他相关变量为协变量，分析了“客观微环境-主观环境感知-居民安全感知”之间的影响机制。研究得出，① 街景绿视率对居民安全感知的作用受到建成环境失序因素的中介。即，绿视率上升导致建成环境失序的降低，进而导致居民不安全感的下降。② 街景绿视率对居民安全感知的作用受到直接受害经历因素调节。即，有直接受害经历的个体，比没有直接受害经历的个体，绿视率对其居民安全感知的影响作用更大。③ 居民安全知感受到街景路灯比例和直接受害经历因素的交互作用影响。即，街景路灯比例对居民安全感知没有显著的直接作用，但是它通过直接受害经历间接影响居民的安全感知水平。④ 街景绿视率对居民安全感知的作用受到建成环境失序中介，同时这个作用受非正式社会融入调节。因此，间接效应分析的结果表明，街景微环境对居民安全感知的影响存在中介效应、调节效应、交互效应。

(4) 个体安全感知的预测及社区安全感知空间模拟的实现。基于街景微环境的直接作用和间接作用结果分析，居民安全感知的形成是由个体因素、客观环境因素、主观环境因素共同驱动的复杂机制。因此，本研究以二分居民安全感知值为结果变量，分别选取二项式 logistics 线性模型，随机森林和支持向量机两种非线性模型，对个体安全感知进行模拟。选取最优的模型，设计一套居民安全感知预测方法，以实现对研究区安全感知的空间模拟。研究得出，① 非线性模型比线性模型的模拟效果好，基于随机森林的安全感知预测模型是最优模型，精度达到 0.7。② 安全感知的空间模拟结果良好，验证精度达  $MSE=0.037$ ，本研究探索出了一套效果良好的利用小样本问卷调查实现大范围安全感知空间模拟的方法。即，按照六普统计年鉴数据，对研究区所有个体的个体属性进行模拟，再耦

## 摘要

---

合所属社区数据，模拟出研究区所有个体安全感知相关属性数据。再利用随机森林训练好的模型，对所有个体的安全感进行预测，最后以均值的方式聚合到社区尺度，实现社区安全感的模拟，并完成预测效果的检验。③ 荔湾区、海珠区和白云区是不安全感最高的区域，而黄埔区、番禺区、天河区是安全感最高的区域。这说明安全感知的空间分布可能与区域人口结构、环境失序水平关系较大，而与经济发展水平关系较小。

本研究的结论可以为城市规划和城市管理部门的城市建成环境治理提供参考依据，例如管理部门可以快速探测出社区环境较差，或安全感评分较低的区域，从而采取对应措施治理城市环境，提升城市品质。本研究的一套社区安全感知模拟方法，可以帮助警务部门实现基于大数据的居民安全感知调查，服务于警务部署和社会治安防控。

**关键词：**安全感知；街景微环境；开放数据；影响因素；机制分析；空间模拟；预测；制图

## 摘要

---

## Abstract

Safety perception, which is also translated as perception of security or fear of crime, is closely related to residents' daily life and affects individual physical and mental health and behavior. In recent years, Chinese government has emphasized that effective social governance and good public order should be developed, and a strong sense of fulfillment, happiness, and security should be improved. Safety perception is influenced by individual and environmental factors simultaneously. Therefore, exploring the relationship between environment and safety perception and conducting the spatial simulation of community-level safety perception will not only enrich the theoretical research on safety perception but also guide the programs of social governance. Previous research on the relationship between the environment and safety perception has certain deficiencies, including the following aspects: 1) Traditional analysis of environmental factors related to safety perception lacks attention to objective microenvironmental factors; 2) Traditional analysis of the relationship between objective environment and safety perception lacks attention to indirect effects of objective environment; 3) The traditional measurement of safety perception and its related environmental factors lack a set of objective, real-time, and low-cost measurement approaches.

Hence, taking seven central urban areas of Guangzhou (Yuexiu District, Liwan District, Haizhu District, Huangpu District, Tianhe District, Baiyun District, and Panyu District) as the research area, using Baidu street-view data, criminal reports data, questionnaire data, Census data, POI data, mobile phone data, and satellite imagery data, this study explores the relationship between environment and safety perception and conducts a spatial simulation task of the community-level safety perception. The main research results are as follows:

(1) Extraction of micro-environment elements of street-view images. Based on Baidu Street View Map, the microenvironment element extraction algorithm flow of “data acquisition-element extraction-spatial analysis” was proposed. Using internet data capture technology, the street-view images with geographic coordinates in the research area of Guangzhou were collected. With the deep learning technology, this research conducted three tasks of street-view image element algorithm. Using the

## Abstract

---

platform of the National Supercomputing Center of Sun Yat-sen University, the feature extraction of all street-view pictures was completed. The microenvironment results were verified by satellite images. The study found that: ① A street-view microenvironment incivilities index system can be established through deep learning technologies. ② In terms of spatial distribution characteristics, factors such as wealth, beauty and green view index are distributed in the main urban areas and areas with high greenery levels. The proportion of buildings in the main urban area is relatively high, and the level of sky openness is low. The proportion of buildings in the outer urban area is low, and the sky openness level is high. ③ In terms of spatial agglomeration characteristics, factors such as wealth, beauty, and green view index are highly correlated, and high-value areas are mainly distributed in the main urban area, Ersha Island, parks, and university towns. The agglomeration characteristics of building proportions and sky openness are obvious and negatively correlated.

(2) Analysis of the direct effect of the street-view microenvironment on residents' safety perception. Taking individual safety perception as the dependent variable, a series of individual attribute indicators, community structural characteristics, community environmental characteristics, and community microenvironment characteristics as independent variables, and this study uses OLS multiple linear regression model to conduct a series of model analyses. The study found that: ① Compared with the street-view wealth and beauty index, the green view index is an ideal indicator of the neighborhood disorder and is a microenvironment indicator that significantly affects the safety perception; ② the model of adding a series of microenvironment elements is better than the model without microenvironment elements, indicating that the addition of streetscape microenvironment elements is beneficial to optimize the safety perception model; ③ Gender, perceived physical disorder and social disorder, social integration, proportion of community commercial facilities, and proportion of community governmental institutions are factors that significantly affect the safety perception.

(3) Analysis of the indirect effect of street-view microenvironment on safety perception. This study uses residents' safety perception as the dependent variable, specific street-view microenvironment as the independent variable, direct victimization experience and informal social integration as the moderating variables, physical disorder as the mediating variable, and other related variables as covariates. Then, this

## Abstract

---

study analyzes the impact mechanism between objective microenvironment, subjective environmental perception, and safety perception. The research results show that: ① the effect of street-view green index on the safety perception is mediated by the physical disorder. That is, the increase in the street-view green index leads to a decrease in the physical disorder, which in turn leads to a decrease in the safety perception of residents. ② The effect of street-view green index on the safety perception is moderated by direct victimization experience. That is, the street-view green index weakens the impact of direct victimization experience on safety perception. ③ Residents' safety perception is influenced by the interaction of the street-view streetlight percentage and direct victimization experience. That is, community streetlight percentage has no significant direct effect on safety perception, but a higher percentage of community streetlight can promote the decline of victimized residents' fear of crime compared with those without victimization experience. ④ The effect of street-view green index on the fear of crime is mediated by the physical disorder, and this effect is moderated by informal social integration. Hence, the results of indirect effect analysis show that the impact of street-view microenvironment factors on resident's safety perception has mediation effects, moderation effects and interaction effects.

(4) The prediction of individual safety perception and the spatial simulation of community-level safety perception. Based on the analysis of the direct and indirect effects of street-view microenvironment, the formation of residents' safety perception is a complex mechanism driven by individual factors, objective environmental factors, and subjective environmental factors. Hence, taking the binary residents' safety perception as the dependent variable, this study selects a linear model (binomial logistics model) and two non-linear models (random forest and support vector machine) to simulate the individual safety perception, separately. Then, the optimal model is selected, and an approach of large-scale community-level safety perception is designed, and community-level safety perception of the study area is predicted. The study found that: ① the simulation effect of the nonlinear model is better than the linear model, and the prediction model of safety perception based on the random forest is the best model, with an accuracy of 0.7; ② The spatial simulation result of safety perception is good, with the verification accuracy of MSE=0.037, which indicates that this study has achieved a method of using a small sample of questionnaires to modeling a large-scale spatial distribution of the safety perception. Specifically, this study simulated the

## Abstract

---

individual attributes of all individuals in the study area based on the data from the Sixth China Census. Then, the community data was assigned to the individuals, and all the relevant safety perception attribute data of all individuals in the study area were obtained. The model trained in the random forest was used to predict the safety perception of all individuals and finally aggregated to the community scale by average value to realize the simulation of the community-level safety perception. ③ Liwan District, Haizhu District, and Baiyun District are the areas with the highest sense of insecurity, while Huangpu District, Panyu District, and Tianhe District are the areas with the highest sense of safety. This shows that the spatial distribution of safety perception may have a greater relationship with the community-level population structure and environmental conditions, but a smaller relationship with the distribution of the economic level.

The conclusions of this study can guide urban planning and urban management departments to improve the urban built environment. For example, government apartments can find areas with disorderly neighborhood environments or low safety perception scores, so as to take corresponding measures to improve the urban environment and quality. The simulation method of community-level safety perception in this study can help police departments realize the survey of large-scale safety perception based on big data, and serve the police deployment and social security control.

**Key words:** Safety Perception; Street-view Microenvironment; Open Data; Influencing Factors; Mechanism Analysis; Spatial Simulation; Prediction; Mapping

## 目录

摘要 .....	I
Abstract.....	V
目录 .....	IX
图目录 .....	XIII
表目录 .....	XV
第一章 绪论.....	1
1.1 研究背景和意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	3
1.2. 概念界定与基础理论.....	4
1.2.1 居民安全感知概念界定.....	4
1.2.2 环境设计预防犯罪理论.....	5
1.2.3 破窗理论.....	7
1.3 环境因素的测量.....	8
1.3.1 建成环境要素的测量.....	8
1.3.2 社会环境要素的测量.....	9
1.3.3 环境失序要素的测量.....	10
1.4 安全感知的影响因素.....	12
1.5 安全感知的测量与模拟.....	15
1.5.1 安全感知的测量.....	15
1.5.2 基于时空视角的安全感知测量.....	18
1.5.3 安全感知的空间模拟.....	19
1.6 国内外研究述评.....	20
第二章 研究问题与研究设计 .....	23
2.1 研究问题.....	23
2.2 研究区域.....	25
2.3 研究内容及总体技术路线.....	26
2.3.1 研究内容.....	26
2.3.2 总体技术路线.....	27
2.3.3 论文框架.....	29
2.4 研究数据.....	31
2.4.1 广州市路网数据.....	31
2.4.2 百度街景图像数据.....	31
2.4.3 问卷调查数据 .....	35
2.4.4 六普人口统计数据 .....	37
2.4.5 POI 设施点数据 .....	38
2.4.6 遥感影像数据 .....	38
2.4.7 手机基站数据 .....	39

## 目录

---

2.4.8 犯罪案件数据.....	40
2.5 小结.....	41
第三章 基于街景图像的安全感知相关微环境要素提取.....	43
3.1 引言.....	43
3.2 基于深度学习技术的微环境要素提取模型实验建模.....	45
3.2.1 广州超算云计算平台.....	45
3.2.2 基于 SE-Inception 卷积神经网络的感知环境要素模型构建.....	45
3.2.3 基于 FCN-8s 全卷积神经网络的客观环境要素分割模型构建.....	50
3.2.4 基于 Faster-R-CNN 卷积神经网络的垃圾目标检测算法.....	53
3.3 广州市中心城区环境要素的描述统计及可视化分析.....	57
3.3.1 感知微环境要素图像评价结果.....	57
3.3.2 客观微环境要素图像分割结果.....	61
3.3.3 垃圾微环境要素目标检测结果.....	66
3.3.4 相关分析与结果验证.....	67
3.4 街景微环境空间集聚特征分析.....	69
3.5 小结.....	70
第四章 微环境要素对安全感知的直接作用分析.....	73
4.1 引言.....	73
4.2 研究方法.....	74
4.2.1 OLS 多元线性回归模型 .....	74
4.2.2 模型评价指标.....	76
4.2.3 变量设定.....	77
4.3 研究结果.....	83
4.3.1 描述统计结果.....	83
4.3.2 微环境对安全感知的直接作用分析.....	84
4.3.3 环境失序程度的最佳微环境衡量指标.....	85
4.4 小结.....	86
第五章 微环境要素对安全感知的间接作用分析.....	89
5.1 引言.....	89
5.2 研究假设.....	91
5.2.1 街景绿视率对安全感知的作用受建成环境失序因素中介 .....	91
5.2.2 街景绿视率对安全感知的作用受直接受害经历因素调节 .....	92
5.2.3 街景路灯比例和直接受害经历对安全感知的作用存在交互效应 .....	92
5.2.4 街景绿视率对安全感知的作用受建成环境失序中介，同时这个关系受非正式社会融入调节.....	92
5.3 研究方法.....	93
5.3.1 中介效应分析.....	93
5.3.2 调节效应和交互效应分析 .....	95
5.3.3 有调节的中介效应分析.....	96
5.3.4 变量设定.....	97
5.4 研究结果.....	98
5.4.1 中介效应结果.....	98
5.4.2 调节效应和交互效应结果.....	99

## 目录

---

5.4.3 有调节的中介效应结果 .....	101
5.5 小结 .....	103
第六章 居民安全感知预测及社区安全感知模拟 .....	107
6.1 引言 .....	107
6.2 研究方法 .....	110
6.2.1 线性和非线性回归模型设定 .....	110
6.2.2 评价指标 .....	114
6.2.3 变量设定 .....	116
6.2.4 研究单元 .....	116
6.3 实验结果与分析 .....	119
6.3.1 基于手机数据的个体安全感知预测模型的输入变量修正 .....	119
6.3.2 机器学习的数据预处理 .....	121
6.3.3 基于二项式 logistics 回归模型的安全感知预测结果 .....	122
6.3.4 基于 RF 随机森林模型的安全感知预测结果 .....	124
6.3.5 基于 SVM 支持向量机模型的安全感知预测结果 .....	127
6.3.6 安全感知分类的各算法对比分析 .....	130
6.4 社区安全感知预测与验证 .....	130
6.4.1 研究区所有个体属性数据生成及预测输入数据确定 .....	131
6.4.2 研究区个体安全感知值预测与验证 .....	133
6.4.3 社区安全感知空间化结果 .....	134
6.5 小结 .....	135
第七章 总结与展望 .....	137
7.1 研究结论 .....	137
7.2 论文创新点 .....	139
7.3 不足之处与展望 .....	140
参考文献 .....	143
附录 .....	157
附表 1 点状尺度的客观街景微环境空间分布 .....	157
附表 2 基于调研社区的预测值和真实值不安全感对照表 .....	160
博士期间发表的学术论文、参与的科研项目和获得的奖励 .....	163
致谢 .....	165

## 目录

---

# 图目录

图 2-1 研究问题的内在逻辑 .....	25
图 2-2 技术路线图 .....	28
图 2-3 论文框架流程图 .....	30
图 2-4 研究区路网地图 .....	31
图 2-5 该地址街景照片 .....	31
图 2-6 广州市研究区街景采样点 .....	31
图 2-7 不同 heading 参数设置示意图 .....	31
图 2-8 Pitch 参数设置示意图 .....	31
图 2-9 社区安全感知问卷 .....	31
图 2-10 研究区六普人口普查行政单元示意图 .....	31
图 2-11 POI 设施点分布 .....	31
图 2-12 大气校正后的遥感影像数据 .....	31
图 2-13 基于泰森多边形的手机基站覆盖范围分布 .....	40
图 2-14 基于核密度的犯罪案件空间分布 .....	41
图 3-1 微环境要素提取流程图 .....	44
图 3-2 MIT place pulse 数据收集平台 .....	44
图 3-3 MIT place pulse 数据集样本及其感知维度评分示例 .....	44
图 3-4 基于 SE-Inception 的 CNN 模型结构图（改绘自 Hu, Shen & Sun, 2018） .....	44
图 3-5 模型预测结果示意图 .....	44
图 3-6 FCN 网络结构图（改绘自 Long et al., 2015） .....	50
图 3-7 上采样过程图（改绘自 Long et al., 2015） .....	51
图 3-8 全卷积网络工作流程图 .....	53
图 3-9 带有垃圾的街景照片标注示例 .....	54
图 3-10 Faster-R-CNN 基本框架（自绘，改自 Ren et al., 2015 & 2016） .....	55
图 3-11 Faster-R-CNN 图像垃圾检测示例 .....	57
图 3-12 感知微环境要素频率分布直方图 .....	58
图 3-13 样点尺度下感知微环境要素空间分布结果 .....	59
图 3-14 社区尺度下感知微环境要素空间分布结果 .....	60
图 3-15 社区尺度下绿视率和水视率微环境要素空间分布结果 .....	62
图 3-16 社区尺度下天空开阔程度和建筑物比例微环境要素空间分布结果 .....	62
图 3-17 社区尺度下路灯比例和多样性微环境要素空间分布结果 .....	62
图 3-18 客观微环境要素频率分布直方图 .....	62
图 3-19 垃圾要素空间分布结果 .....	66
图 3-20 社区尺度的垃圾要素空间分布结果 .....	66
图 3-21 垃圾要素频率分布直方图 .....	67
图 3-22 基于遥感影像的绿视率验证分析 .....	66
图 3-23 基于遥感影像的水视率验证分析 .....	66
图 4-1 微环境要素直接作用分析流程图 .....	74
图 4-2 最小二乘法计算原理 .....	75
图 4-3 样本社区分布 .....	77
图 4-4 变量相关分析图 .....	82

## 图目录

---

图 5-1 微环境要素间接作用分析流程图.....	91
图 5-2 中介效应示意图 .....	94
图 5-3 调节变量示意图 .....	96
图 5-4 有调节的中介效应模型示例 .....	97
图 5-5 中介效应可视化结果 .....	99
图 5-6 调节效应图 .....	101
图 5-7 有调节的中介效应图 .....	102
图 6-1 社区安全感知模拟流程图 .....	109
图 6-2 随机森林分类器图 .....	112
图 6-3 SVM 超平面划分 .....	114
图 6-4 输入变量的社区环境指标 .....	118
图 6-5 不同数据源的受教育水平和外来人口比例指标对比 .....	118
图 6-6 不同数据源的人口密度指标对比.....	118
图 6-7 二项式 logistics 模型 ROC 曲线结果 .....	124
图 6-8 不同树选择下的分类精度 .....	125
图 6-9 重要性程度排序 .....	127
图 6-10 随机森林模型 ROC 曲线结果 .....	127
图 6-11 支持向量机模型 ROC 曲线结果 .....	129
图 6-12 真实值与预测值散点图比较 .....	133
图 6-13 广州市研究区不安全感空间分布.....	135

# 表目录

表 1-1 安全感知测量题项 .....	16
表 1-2 安全感知的不同测度工具 .....	17
表 2-1 研究问题总结 .....	25
表 2-2 不同街景镜头参数设置详细介绍.....	25
表 3-1 垃圾目标检测实验结果 .....	56
表 3-2 不同行政区域感知微环境要素结果统计.....	58
表 3-3 不同行政区域客观微环境要素结果统计.....	61
表 3-4 不同行政区域客观建筑垃圾要素结果统计.....	67
表 3-5 相关分析结果表 .....	68
表 3-6 全局空间自相关 Moran's I 值 .....	70
表 4-1 样本变量描述统计分析 .....	83
表 4-2 微环境对不安全感知的直接作用分析.....	84
表 4-3 环境失序程度的最佳微环境衡量指标分析.....	86
表 5-1 调节效应选取规则 .....	96
表 5-2 建成环境失序在街景绿视率-不安全感关系中的中介效应分析 .....	99
表 5-3 调节效应和交互效应检验结果 .....	100
表 5-4 社会融入调节街景绿视率-建成环境失序-不安全感中介效应模型回归结果 .....	103
表 6-1 混淆矩阵原理 .....	115
表 6-2 基于混淆矩阵的常用评价指标 .....	115
表 6-3 真实值和预测值的评价指标 .....	115
表 6-4 个体安全感知模型的输入变量 .....	117
表 6-5 基于二项式 logistics 回归的模型对比 .....	121
表 6-6 二项式 logistics 回归模型估计 .....	122
表 6-7 二项式 logistics 回归模型逐步回归 .....	123
表 6-8 卡方检验结果 .....	123
表 6-9 二项式 logist i cs 回归模型混淆矩阵结果 .....	124
表 6-10 混淆矩阵 .....	126
表 6-11 初步 SVM 模型混淆矩阵结果 .....	128
表 6-12 SVM 模型调参汇总结果 .....	129
表 6-13 最优 SVM 模型混淆矩阵结果 .....	129
表 6-14 各算法不同分类安全感知评价精度指标.....	130
表 6-15 模拟数据分布表 .....	132
表 6-16 模拟数据分布表指标解释 .....	132
表 6-17 基于样本社区的验证结果指标 .....	133
表 6-18 行政区尺度安全感知预测结果 .....	134

## 表目录

---

# 第一章 绪论

## 1.1 研究背景和意义

### 1.1.1 研究背景

安全感知，与居民日常生活息息相关，是一个重要的社会话题。近年来，中央政府不断强调要完善社会治理，建设和谐社会，提升人民群众的幸福感、获得感、安全感。国内外研究已经表明，安全感能影响个体的身体健康（Stafford, Chandola, & Marmot, 2007; Macassa et al., 2017）、心理健康(Whitley & Prince, 2005; Alfaro-Beracoechea et al., 2018）、日常活动 (Stodolska et al., 2013)等多方面。

近年来，全国各地各级部门广泛开展了群众安全感的调查。安全感调查主要有两方面的目的，一方面是识别出城市内部安全感较低的社区或街道，从而进一步采取措施提升居民安全感；另一方面，不同于客观的犯罪率，群众安全感是真实直接的居民主观感受，可以作为对各级部门治安工作成效评价的一个重要指标。但是，我国当今的城市规划建设，一系列硬件环境设施等都已经完善，甚至走在世界前列，但是针对居民的主观感知的规划建设还很少。比如，建设健康城市、平安城市。这种情况下，要提升国家软实力，居民安全感的研究和相关城市规划开展十分必要。

当今，大数据时代来临，通过大数据推动社区治理创新，可以提升社会治理的高效性和准确性。2017年12月8日，中共中央政治局集中学习了国家大数据的战略，总书记强调应该结合大数据实现现代化建设，用大数据服务于社会治理。同时，近十年来，一系列开放数据源（如手机数据、街景数据、遥感数据、POI设施点数据、微博等社交媒体数据）涌现。这些开放数据源具有数据量大、数据更新快、内容丰富、有地理位置信息等特征，是属性丰富的地理信息数据。这些数据驱动着城市研究向精细化发展，越来越多基于开放数据的个体行为、活动路径研究开展（敬峰瑞，孙虎 & 龙冬平, 2017 a&b; Song et al., 2019; 姚海芳, 冯天楠 & 刘劲松, 2020）。

街景图像数据作为一种可获取性强的开放数据源，已经被广泛应用于科学研  
究领域（Rundle et al., 2011; Kelly et al., 2012; Li et al., 2015; Helbich et al., 2019;

Jing et al., 2021)。相关研究已经表明, 街景图像可以测量出社区的建成环境状态, 即识别出社区环境是否干净、整洁、有序等 (Rundle et al., 2011); 以及社区的自然环境状态, 如绿视率等 (Li et al., 2015)、天空开阔程度等 (Zeng et al., 2018)。此外, 街景图像也可以反映出该社区的人口统计特征、社会经济状况等 (Gebru et al., 2017)。借助街景图像数据源, 可以相对客观、快速地测量出城市各区域的社区环境状况。

然而, 结合开放数据开展的环境感知研究还相对缺乏, 特别是结合开放数据和居民安全感知的相关研究。开放数据作为一类高时空分辨率的数据, 可以更加精细地刻画个体的环境暴露状态, 可以视为影响居民安全感知的潜在微环境变量。个体的日常活动一般基于道路展开, 而基于街景图像的微环境指标也是基于道路拍摄并提取获得。因此, 基于街景图像的微环境指标可能是影响居民安全感的重要因素。基于街景图像, 可以帮助更全面地解释出居民的社区安全感知的影响机制。

深度学习是人工智能研究最为热门的一个领域, 结合深度学习的计算机视觉研究发展迅速。在有监督学习领域, 以图像分类、图像目标探测、图像分割、图像标注、图像生成领域, 深度学习都产生了变革性的影响, 将传统算法的精度大为提高, 而且新的深度学习算法更新迭代速度很快。在街景图像研究领域, 基于街景图像的图像分类(如基于 CNN 卷积神经网络的道路交通护栏分类(杜坤等, 2018))、图像分割 (如基于 FCN/PSPNet 深度学习算法的图像分割 (Yao et al., 2019))、目标检测 (如街道中的字体检测, 王娅君 (2017)) 也不断有研究深入分析。因此, 街景图像属性丰富, 包含丰富的道路景观特征, 运用深度学习方法, 可以挖掘出城市环境中的多维度微环境特征, 从而可以为与空间行为有关的城市研究提供重要的数据源, 从而开展更加微观尺度的环境和感知研究, 发掘更多环境与感知之间的关系和机制。在居民安全感知研究领域, 街景图像更是大有益处。

在 GIS 空间制图研究领域, 结合环境感知的研究开展的相对较少。传统的安全感测量一般以社区、街道、或区为单位, 通过问卷调查或电话访问的形式开展。问卷调查的方式虽然可以直接获取真实的居民安全感知数据, 但是成本较高、主观性较强、耗时耗力, 不适合大规模开展。当今, 基于多源开放数据, 对个体安全感进行模拟, 可以实现高时间和空间粒度的社区安全感知度量。

因此，借助开放数据，分析居民安全感知的微环境影响因素，进而开展空间格局模拟，获得社区安全感的空间分布状况，可以帮助政府开展针对性工作，提升社区居民安全感水平，提升社区治安安全状况，保证社会稳定、长治久安。

## 1.1.2 研究意义

本研究具有理论意义和实践意义。

### (1) 理论意义

环境因素已经被证明是影响居民安全感知的重要因素，包括邻里的建成环境、建成环境感知状态、自然环境特征、自然环境感知状态等。本研究主要的理论贡献有若干点。

第一，一些客观的环境要素大规模测量一直是犯罪、健康等城市问题研究领域难以突破的问题，相关研究还主要集中在基于问卷调查获取感知环境失序，以及基于街景图像进行小规模的客观环境评价。利用街景图像数据，本研究精细化、大规模测度各项微环境要素，据此探讨微环境因素对于安全感的影响机制。这可以丰富破窗理论（Broken Windows）和环境设计预防犯罪（CPTED）等环境设计理论在居民安全感知研究领域的发展。

第二，本研究尝试加入微环境因素来解释居民安全感知，以及考虑微环境因素对于安全感知的直接和间接效应，从而探测出与居民安全感知相关的微环境要素。因此，本研究有利于改进传统的居民安全感知模型，提供模型的预测能力。

第三，传统的安全感知模型都是线性模型，但是实际的场景中并非如此。个体的不安全感可能开始时会随着环境刺激而增加，但是达到一定限度后可能就不会明显增加，这与心理学中的脱敏效应类似。本研究尝试比较线性模型和非线性模型对于居民安全感知的解释能力，检验非线性模型是否能够更准确地解释居民安全感知。从而为以后的安全感知模型建模与预测提供理论方法上的指导。

第四，关于环境和安全感知的研究，西方已经有一系列理论模型，包括受害模型、失序模型、社会融入模型等。然而，这些理论模型是否在中国背景下的安全感知研究同样适用，还有待检验。本研究以中国城市为研究区，验证西方安全感理论在不同文化背景下的适用性，从而丰富相关犯罪学理论在安全感知研究领域的发展。

### (2) 实践意义

长期以来，安全感及其环境要素的调查主要是通过随机抽样的问卷调查方式所得。调查时间长、人力、物力成本高，难以开展大规模、多时相的调查。通过引入街景图像数据，与社会经济数据和问卷调查数据相耦合，建立居民安全感知预测模型，为大规模、长期、动态的安全感评估提供一种客观高效的技术方法，分析结果可以为社会治安防控提供重要指导。具体来看，本研究的实践意义主要有以下三点。

第一，广州市作为我国典型的大城市之一，街区环境复杂多样。本文首次利用街景数据，针对广州市这一特定城市，克服之前研究方法的弱点，建立一套基于街景图像的环境失序评价体系，并且精细化刻画广州市 7 个中心城区的环境及环境失序空间格局；

第二，本研究通过分析客观环境对居民安全感知的作用及其影响机制，从而可以对城市规划和城市管理提供实际指导和建议。政府部门通过环境设计等一系列针对性措施，可以提升区域内居民的安全感水平。

第三，广州作为我国改革开放的前沿，长久以来，外来人口众多，特别是与广大亚非拉地区联系紧密。近年来，在政府部门的强力治理下，犯罪率已经保持在相对稳定的水平，社会客观治安状况已经得到改善，但是居民的主观安全感知变化情况还缺乏科学测度。而当今社会的城市生活导致居民的生活压力增大，精神健康越来越受到重视，及时获取城市居民的社会治安安全感知状况关系于社会稳定。本文首次利用街景数据，耦合多源数据，针对广州市这一特定城市，克服之前问卷调查安全感的弱点，建立一套基于多源数据的社区安全感测度方法，并且计算广州市社区尺度的安全感水平。

## 1.2. 概念界定与基础理论

### 1.2.1 居民安全感知概念界定

虽然安全感知的研究比较丰富，但是安全感知的定义还没有统一的标准，存在诸如治安安全感、交通安全感、工作安全感、心理安全感等多种分类。本研究的安全感知是狭义上的治安安全感，反映了居民对于生活社区的治安水平的稳定感知状况。

英美国家最早开始安全感知的研究，常用的术语包括 perception of safety 或 fear of crime (被害恐惧感)。Conklin(1971) 定义为个体在社区中的一种对于遭受被害的一种感知状况。Garofalo(1981) 定义为一种有身体伤害的威胁所产生的危险感和焦虑，是一种情绪反应，由环境中与犯罪有关的某些方面的感知线索引发。Ferraro(1987) 认为安全感知是关于被害或者是与犯罪相关的标志的情感上的焦虑或恐惧。综合来看，在理论上，本研究定义安全感知是个体在所处社区中的关于遭遇各类犯罪行为的一种焦虑状态。

居民安全感知(个体尺度): 指个体尺度下的居民对于遭遇犯罪的担忧水平。

居民安全感知（社区尺度）: 指聚合到社区尺度下的该社区所有居民安全感的均值。

### 1.2.2 环境设计预防犯罪理论

环境设计预防犯罪理论 (Crime Prevention Through Environmental Design) 缩写为 CPTED 理论，是一种用于解释环境设计可以增强居民安全感以及提升居民生活质量的跨学科理论。

人类与环境的关系是一个悠久的话题。《圣经·旧约》描述到，Adam 和 Eve 居住在伊甸园，但是却违背规则偷吃禁果，结果被驱离伊甸园。从人地关系来看，这个故事就是由人、环境、犯罪组成。在工业革命之前，自然环境是影响人类活动的关键因素。而工业革命之后，自然环境被人类大力改造，人文环境逐渐成为影响人类活动的主要环境因素。如今，整个世界都处于城市化进程，城市人口已经占全球人口大部分，城市是人地关系最密切的空间。学者也重点关注于城市环境对居民行为和生活方式的影响。

Jacobs (1992) 在她的著作《*Death and Life of Great American Cities*》提出了街道眼理论。她指出，较高密度、混合利用、紧凑布局的建成环境为人群间的彼此保护创造了条件，有助于降低犯罪率和提高安全感；蔓延、单一的城市空间则会成为孕育犯罪的温床。理想的城市空间就是当人们走在陌生的街道上时会感觉安全。在城市规划设计时，街道需要确保是有活力的和多样化的，以提升行人的安全感。针对美国、澳大利亚等地的研究也证实，功能混合、小尺度街区以及可达性良好的人流密集社区普遍拥有较高的居住安全感 (Foster, Giles-Corti &

Knuiman, 2010; Foster et al., 2013)。

她认为，一个城市街道能够把街上的陌生人变成一个安全的资产，作为成功的城市邻里总是那样做的那样，需要包括以下三个主要特征：

- A. 公共和个人的空间应该存在明确的边界。这两类空间不能相互渗透。
- B. 街区道路应该至少有几双眼睛，它们可以被看成这些道路上的自然所有物。街道上的建筑可以容纳陌生人，必须面向街道，可以确保行人和陌生人的安全。
- C. 人行道上必须持续有使用者，增加街道上有效的“眼睛”，沿街建筑要保证有足够的数量上的人观察人行道，因为没有人喜欢坐在门廊上看到窗外空荡荡的街道。

Jacobs 的街道眼理论是 CPTED 思想的萌芽状态。而 CPTED 概念正式的提出是在犯罪学家 Jeffery (1971) 的著作《*Crime Prevention Through Environmental Design*》。基于 Bentham (1983) 的功利主义模型和 Skinner(1965) 的刺激-反应模型，Jeffery 提出了环境设计预防犯罪的学术概念 (Jeffery, 1971)。基于心理学的基本学习理论，他认为环境的刺激形成和控制行为，建议利用环境加强人们所希望的行为，而不是惩罚那些不起作用的行为。他认为环境设计预防犯罪涉及生物学、心理学、犯罪学、城市地理、城市规划等多个领域。

1972 年，建筑学家 Oscar Newman 提出了防卫空间的概念，即通过环境设计控制和减少犯罪 (Newman, 1972)。Newman 的犯罪预防方法包括属地、监视、图像、安全区域的概念。他的方法已经被美国政府、企业、学者广泛采用。二者相比较，Jeffery 的犯罪预防方法更加综合，Newman 偏重于环境设计。

CPTED 自从概念问世以来，不断更新发展，并与其他基于场所的策略（比如情景犯罪预防和环境犯罪学）相结合。CPTED 理论已经在地方和国家层面得到广泛运用。具体看，CPTED 的原则包含六个要素：领域性 (territoriality)、监视 (surveillance)、访问控制 (access control)、目标增强 (target hardening)、场景或维护 (image/maintenance)、事件支持 (activity support)。

领域性是物品的所属情况。如果某物品没有确定的所属者，则该物品被损坏的可能性会变大，并且该物品周围区域受到损坏的可能性同时会上升。如果某物品的所属情况是确定的，则一旦其被损坏，相关负责人将对这种越轨行为采取措

施。因此，有领域性会导致遭遇犯罪的可能性降低。

监视。通常情况下，人口稠密的地方会提供自然的监视功能，从而抑制犯罪活动。因此，对街区道路的监视可以遏制损害活动的出现。

访问控制。访问控制表示，如果进入某些场所的前提是对访客开展一些检查措施，那么这些场所不太可能发生各种损害活动。

目标增强。目标增强指对特定目标进行保护，比如增加栅栏、围墙等措施。

场景或维护。场景是空间和建筑等的外部条件。为了防止犯罪，有必要定期维护场景以防止衰败的现象出现。

事件支持。一些区域的热度本身比较低，周围人口少，缺乏治理，从而会出现失序的特征。对于这种情况，需要人为增加活动和事件以提高监视水平。事件支持就是人为增加该地区的人员活动，以实现对犯罪预防的支持。

### 1.2.3 破窗理论

作为 CPTED 理论的补充，Wilson 和 Kelling 提出了破窗理论 (Wilson & Kelling, 1982)。他们强调社区秩序的维持对于增强社区居民的安全感知很关键。失序和犯罪经常在社区水平上相关联。如果一个建筑物的窗户破坏了而且没有被修复，那么这个建筑物剩下的窗户就会很快被破坏。原因在于，被破坏的窗户表明了这个地方是无人照看的，那么破坏窗户没有成本。这种情况下，就会产生恶性循环，出现越来越多窗户被破坏的结果。破窗事件可能导致社区控制的崩溃。Wilson & Kelling (1982) 因此主张控制低级反社会行为和轻微犯罪事件，防止重大犯罪事件，从而保持整个社区健康。在这样做的过程中，他们说，警察在维护社区秩序方面的作用很重要。警察应该与他们所辖的社区保持联系，保持社区秩序，而不仅仅是预防社区犯罪事件。Sampson & Raudenbush (2004) 拓展了破窗理论，认为垃圾、涂鸦、废弃的汽车等环境都会使犯罪者认为居民对社区漠不关心，导致这个邻里的居民对邻里犯罪现象缺少协作。因此，犯罪者容易在该社区实施犯罪。

基于破窗理论，社会融入理论相应产生。社会融入理论认为高水平的非正式社会融入可以减少失序感知程度。社会融入被定义为个体对于所在社区的归属感以及依恋感。一个高信任以及团结水平的社区，会增强居民的期待，认为这个社

区的社交网络和纽带可以抑制犯罪活动。

对比 CPTED 理论和破窗理论，二者相互补充。破窗理论回答了越轨行为发生的原因，指出外在环境会诱发越轨行为，环境是越轨行为的影响因素。CPTED 理论则认为环境设计方案能够阻止越轨行为，良好的邻里环境可以阻止损害活动的出现。

## 1.3 环境因素的测量

### 1.3.1 建成环境要素的测量

在犯罪学和公共卫生领域，一系列文献指出邻里环境中如住宅类型、植被覆盖率可以有效提升居民的身心健康水平(Astell-Burt, Mitchell, & Hartig, 2014)、安全感水平 (Maas et al., 2009)，以及降低区域的犯罪率 (Du & Law, 2016)。建成环境是影响居民社区感知的重要因素(Lee et al., 2008)。

建成环境指标主要包括植被、天空、建筑物等城市客观环境。对于小范围的建成环境测量，主要以使用传统测绘仪器的摄影测量方式获取为主。对于大氛围的建成环境调查，遥感技术以大面积覆盖、时效性高、成本合理等优点，目前广泛运用在城市绿地调查、生态景观格局过程分析等领域。随着遥感影像空间分辨率的提升，高分辨率遥感数据也不断运用于城市环境和居民感知关系的研究中。通过高分遥感数据，可以更精细地提取不同种类植被类型、用地类型，以及提取空间分辨率更高的建成环境信息 (Lee et al., 2008)。

近年来，以 LiDAR (激光雷达遥感技术) 和无人机为代表的新技术以时效性高、精度准确、性价比好的优点，在城市建成环境信息提取领域得到快速应用。LiDAR 高精度的立体扫描技术，可以开展具有三维空间结构且几何体形态不规则的信息获取。如目前有一系列研究利用 LiDAR 提取三维绿量，更加全面的度量了城市绿地功能 (韦雪花 et al., 2013; Yao, Krzystek & Heurich, 2012)。此外，无人机以其环境适应性强的优势，在不同测绘任务中逐渐广泛应用，如林业资源调查、地质调查、工程测量领域 (胡健波 & 张健, 2018)。

大数据时代到来，各领域的开放数据也逐渐成为各类研究的重要数据来源。以 Google 街景、百度街景、腾讯街景为代表的街景数据，由于获取性强、图像

信息全等优点，利用街景图像自动化大批量的提取建成环境要素的研究涌现 (Li et al., 2015; 崔喆, 何明怡 & 陆明, 2018; Gong et al., 2018; 张永霖 & 付晓, 2020)。以街景图片的绿视率为例，“绿视率”最初由青木陽二(1987)使用，指人眼正常视野范畴内观察到的绿化面积的比例。传统上，利用人工拍摄的照片，根据色彩 RBG 分段获取绿色占比的部分，作为绿视率的测度值 (Yang et al., 2009; Jiang et al., 2017; 崔喆, 何明怡 & 陆明 2018)。但是这种方法规模小、更新慢。Li et al. (2015) 基于 Google 街景图像，提出了一种 MGVI 的植被指数，评估了纽约市曼哈顿地区 East Village 的街道的绿视率水平，对城市规划和管理提供了建议。王昱 (2016) 结合街景数据和 LiDAR 数据两种数据源，获取了包括多维绿量、绿色视觉指数、色彩指数在内的多种指标，构建了城市街道美景度评价体系。这些研究大多是基于色彩区间提取法进行植被要素的提取，但是这类技术对于多类、多色绿化要素的整体提取还有局限，还是不能适用于大规模的准确提取和研究。比如，区分绿色时，难以区分自然还是人工的绿色。

与此同时，随着深度学习技术在计算机图像处理和视觉领域的发展，一系列其他建成环境要素的探测和识别方法发展起来。这主要分为两部分，包括图像语义分割技术和图像目标检测技术。传统的图像分割技术，通过逐像素的分类来识别植被、天空、建筑物等。而基于深度学习的图像语义分割技术，通过样本的标准和训练，实现照片的自动分类，可以更加准确地检测出街景图片中不同物体的类别。比如， Badrinarayanan, Kendall & Cipolla. (2017)利用机器学习算法的卷积神经网络工具(SegNet)，提取室内和室外环境语义要素，其效果优于 FCN 模型。Yao et al. (2019)基于武汉市中心城区的街景，利用 FCN-8s 模型刻画了 150 种街景客观环境要素。而目标检测技术如 Faster-R-CNN，可以识别出图片中的特定建成环境的，如各类机动车如自行车、小汽车、公交车等(Espinosa, Velastin & Branch, 2017)。

## 1.3.2 社会环境要素的测量

社会环境主要包括人口密度、GDP 水平等社会结构指标。通常情况下，这些指标的获取来源于人口普查数据和统计数据，然后计算各类指标获取。比如，当前我国主要的统计调查包括 10 年一次的全国人口普查、5 年一次的 1% 人口调

查，以及每年一次的 1‰人口调查，和 GDP 水平指标的调查。然而，这些社会环境动态监测方法具有以下缺点：难以统一调查范式、空间分辨率低、时效性差等，不能满足实时人口动态监测的需要。

信息技术和 Internet 的发展，大数据逐渐构成社会环境动态监测的数据源，一系列新方法和新技术被用来测量社会环境。相关数据从传统统计数据源（如人口普查、问卷调查、田野访谈等）转向基于地理位置的数据，如遥感影像数据（冯甜甜 & 龚健雅, 2010；吴健生 et al., 2015）、网络数据（刘云舒、赵鹏军 & 梁进社, 2018；赵鹏军 & 曹毓书, 2018）、手机信令数据等（王德, 朱查松 & 谢栋灿, 2016）。运用遥感数据产品，如夜间灯光数据可以预测区域贫困水平（Jean et al., 2016；李宗光 et al., 2016），以及区域 GDP 水平（韩向娣 et al., 2012）。综合统计人口数据、基于地理位置的人口数据和人口分布的影响因素，可以实现统计人口的高时空分辨率制图（廖顺宝 & 孙九林, 2010；Stevens et al., 2015）。作为一种重要的基于地理位置（LBS）的数据，基于手机用户的时空分布和旅行轨迹，可以实现对城市空间功能区的识别（钮心毅, 王垚&丁亮, 2017）、建成环境的评估（王德 et al., 2015）、居民出行方式的识别（Wang et al., 2010；Aguilera et al., 2014）、周围人口密度和人口活动的动态监测（钟炜菁 et al., 2017；He et al., 2020）等。

近年来，结合街景图像和计算机深度学习技术，不断有研究开展各种社会环境指标的识别。基于大规模的街景照片及其评分数据，麻省理工学院（MIT）研究团队利用深度学习算法，探索了使用街景照片进行美丽度、富裕度、活力度、安全度等社会环境相关指标的识别，并大规模计算了美国纽约，以及中国的北京、上海等城市的街景城市风貌特征（Zhang et al., 2018；Zhang et al., 2021）。Gebru et al. (2017) 基于大规模街景图像识别小汽车的品牌，从而推断出社区经济水平状态。因此，社会环境的测量也与街景图像不断深度结合。

### 1.3.3 环境失序要素的测量

安全感的相关研究揭示了邻里环境失序是安全感的一个重要预测因素。同时，邻里环境有序/失序的测量也一直是热门的话题。与安全感测量类似，传统的邻里环境失序特征测量是通过问卷得到（Scarborough et al., 2010；Wyant, 2008）。

常规的测量问题如下：“请对以下社区内部安全情况进行评价打分：社区内的公共设施经常被破坏，社区内常见垃圾、车乱停的现象，社区内常见乱涂乱画或者乱贴广告的现象，社区内环境很嘈杂”。回答选项如下从非常不符合到非常符合，以李克特五点量表的形式测量。然而，传统的建成环境失序特征的测量方法存在主观性强、调研地难以扩大、成本大等缺点。有研究就指出当前的一系列社区失序影响安全感的研究都是以感知的失序特征为指标，而实际上客观的失序特征对安全感的影响机制也十分重要，而且可能有不同的结论 (Roccato, Russo, & Vieno, 2011)，但是目前的相关研究还很少。

一系列利用多源大数据的方法已经应用到社区失序状态的测量中。有研究利用波士顿市 1 亿条 911 数据和 200000 条非紧急服务中的建成环境失序指标，揭示了失序和犯罪主要不是来自于公共的线索，而是来自于社区内的失序状态，这个研究发展和解释了破窗理论 (O'Brien & Sampson, 2015; O'Brien, Sampson, & Winship, 2015)。通过 Twitter 数据中与失序相关的帖子与实际犯罪率之间的分析，Williams, Burnap, & Sloan (2017) 研究发现社交媒体数据或许可以作为提供犯罪问题的一个信息来源，还分析了大数据作用于越轨行为研究的局限性。考虑到主观感知的社区失序数据在解释犯罪问题上的局限性，有研究还通过收集在线问题报告网站的数据，分析了这类众包开源数据如何表征社区失序，以及讨论了这种方法对于推进犯罪和地点感知研究的意义 (Solymosi, Bowsers, & Fujiyama, 2017)。

在公共卫生领域，目前已经有一系列研究通过街景建成环境测量来研究环境对个体日常活动、社区社会经济指标的作用机制 (Charreire et al., 2014; Wilson et al., 2012; Badland et al., 2010)。对于建成环境失序，Mooney et al. (2017) 的研究中，十一个受过训练的评判人对 1826 张 Google 街景图像的 9 个维度进行评分并对费城的失序特征进行了核密度估算，发现失序特征的指标与美国统计数据中失业率和房屋空置率显著相关。同时，他指出利用街景图像可以有效的提取失序特征指标。Quinn et al. (2016) 运用类似的手段，通过人工对纽约市 532 个街区的 Google 街景图像打分，使用克里金插值方法对纽约市的失序特征进行了制图，并且他的研究中也详细介绍了街景打分的一系列标准。近年，关于街景是否能有效测量社区环境和社区失序的研究相继发表 (Mooney et al., 2017)。

随着深度学习技术的发展，基于卷积神经网络的图像分类和回归技术也逐渐

运用到一些大规模感知环境要素失序的识别中。比如，基于街景打分，运用卷积神经网络可以实现街道品质高低的分类（甘欣悦 & 余天唯, 2018; 崔成 et al., 2020），并基于此完成了大规模城市空间品质和非正规性的空间制图。前文提到的 MIT Media Laboratory 开展的感知环境要素识别中，美丽度和富裕度也属于街景失序程度的逆向指标（Zhang et al., 2018）。另外，基于深度学习的图像目标检测技术，可以识别出街道景观中的失序类物体，如各类垃圾等(Mittal et al., 2016; 魏书法 & 程章林, 2017; 张鹏程, 赵齐 & 高泽宇, 2019)。

随着信息技术的发展，多源数据的可获取性越来越强，邻里环境失序的测量已经逐渐从以单一的问卷测量方法为主，到利用社交媒体数据、众包网站数据、政务大数据、街景图像数据等多种数据源开发出不同的测量方式。

## 1.4 安全感知的影响因素

影响安全感的因素主要可以分为个体层面（Individual）的变量和社会情景层面（Social Contextual）的变量。

个体层面来看，性别、年龄、种族、教育水平、之前被害情况、感知的社区警务、媒体使用率等都是与安全感有关的因素。个体维度要素对居住安全感知的影响已经发展出了一系列模型。第一个重要模型是脆弱性模型。脆弱性模型假设客观或感知的个体由于身体或社会因素的遭遇犯罪的脆弱性程度，从而导致恐惧感。身体脆弱性的群体，比如女性和老年人，倾向于有更高不安全感。不少研究已经表明，女性比男性有更强的不安全感(LaGrange & Ferraro, 1989; Britto, Stoddart, & Ugwu, 2018)。相比于青年人，老年人表现出更强的不安全感 (Rand, 2008)。而社会脆弱性的群体，因为他们自身的社会脆弱性特征，更有可能有更高的恐惧感，。非裔美国人比白人表现出更强的不安全感 (May et al., 2015)。正规的受教育水平上升，个体的安全感则显著的降低 (Scarborough et al., 2010)。收入较低的群体，有更高的犯罪被害敏感度 (Cho & Park, 2017)。

与脆弱性模型密切相关的是被害模型。被害模型被理论化为之前的受害经历，特别是感知的受害经历的严重程度，使个体对遭遇犯罪更加敏感，会产生更高的恐惧感和风险感知 (Abbott & McGrath, 2017)。这个发现在西方 (Rühs, Greve & Kappes, 2017; Singer et al., 2019)和非西方文献中(Chan & Zhang, 1999; Lee & Cho,

2018)都得到证实。另外，间接的受害经历也会导致更高的被害恐惧感知。一系列证据表明，安全感知是与一些关于受害经历的二手信息源相关的，比如非正式社会网络，以及社交媒体。社交媒体的运用越频繁，会减少个体的安全感 (Kupchik & Bracy, 2009; Chadee & Chadee, 2019)。而 Hanslmaier (2013)发现媒体的使用率是犯罪率与安全感的调节变量，也就是说，频繁的媒体使用，会导致犯罪率与个体安全感知的关系变密切。此外，其他人也有揭示出被害经历和不安全感的关系可能比较弱(Box, Hale & Andrews, 1988; Tseloni & Zarafonitou, 2008)。这种弱关系的存在可能是因为受害者情感的缓和，记忆的减退，犯罪活动的阻止，以及犯罪事件的合理化认知(Box, Hale & Andrews, 1988)。

安全感知不仅受个体因素影响，还是环境因素影响，包括主观和客观环境 (Lee, Park & Jung, 2016; Lis, Pardela & Iwankowski, 2019; Cho et al., 2019; Tchinda & Kim, 2020)。社会情景层面来看，相关的影响因素为感知的社会失序、感知的自然失序、感知的邻里警力、非正式社会融入程度、邻里依恋程度、邻里控制、邻里犯罪率、邻里建成环境特征、社区经济状态等。这些情景变量主要可以分为社会失序类、社会融入类、社区一般结构指标类。破窗理论假设，社区失序现象，如涂鸦、公共场所酗酒、垃圾乱放、汽车遗弃现象等会导致进一步的社区衰败，会使犯罪者认为这个社区的社区融入度不高，缺少社区凝聚力，这就会促使犯罪者实施犯罪，也会使居民产生更高的不信任感觉和恐惧感(Wilson & Kelling, 1982; Sampson & Raudenbush, 2004)。社区失序和安全感之间显著的关系在一系列中西方文献中被发现(Woldoff, 2006; Zhang et al., 2009; Karakus, McGarrell & Basibuyuk, 2010; Porter, Rader & Cossman, 2012; Helfgott, 2020)。还有一些文献发现，一些因素，比如受害经历(Roccato, Russo & Vieno, 2011)、社区纽带(Ross & Jang, 2000)调节了社会失序和安全感的关系。犯罪率可能也可以作为一个客观的风险感知因素和社区失序的因素(Barni et al., 2016)。区域的高犯罪率表示该区域有较高的实际被害风险。但是很多犯罪事件并不是直接在公共空间中可以被看见。一般而言，邻里的犯罪率增高，该邻里个体的安全感知程度就降低 (Rountree & Land, 1998; Brunton-Simith & Sturgis, 2011; Hanslmier, 2013; Breetzke & Pearson, 2014)。但是也有一些研究没有发现二者之间的显著相关关系 (Wyant, 2008; Alkimim, Clarke, & Oliveira, 2013)。另一个社区失序的指标是社区收入水平，社区收入水平被认

为与居民安全感知相关(Scarborough et al., 2010)。

与失序模型相关，集体效能理论 (Sampson & Raudenbush, 2004) 认为一个社区如果拥有高水平的集体效能，可能导致社会失序水平的下降。集体效能视角强调社区凝聚力的重要性和社区控制，分为非正式社会融入和非正式社会控制两个方面。社会融入被定义为个体对于所属社区的归属感以及依恋感 (Adams & Serpe, 2000)。一个高水平信任和团结的社区，会加强居民的期待，认为该社区的网络和纽带可以抑制犯罪活动。相应地，与社区邻居熟悉的居民可能经历更低的不安全感。大多数研究认为社会融入会抑制不安全感的水平(Gibson et al., 2002; Franklin, Franklin & Fearn, 2008)，但是一些文献也证明出社会融入可能导致不一致的发现(Ross & Jang, 2000)。非正式的社会控制，作为社区社会融入的结果 (Wikström & Dolmén, 2001)，代表着实现共同社区观念的社会秩序的维护水平，比如生活在一个安全有序的环境的目标(Sampson & Raudenbush, 1997)。如果个体与周围的邻居建立了积极的社交网络，他们倾向于采取措施阻止可能发生在社区的暴力。一些文献揭示了非正式社会控制和减少的不安全感之间的关系(Yuan & McNeeley, 2017)。

此外，社区警力模型强调公众对警力的感知水平与个体安全感知的水平是相关的(Roh & Oliver, 2005; Scarborough et al., 2010; Maguire et al., 2017)。邻里警力提升了警察在居民中存在的水平，加强了警察和市民之间的实际联系，经常导致居民会对警察有更高的满意度，也会消减他们的不安全感。Scheider, Rowell, & Bezdikian (2003) 指出对警察的满意度和个体的安全感是正相关的关系。Reisig & Park (2000) 认为如果感知的安全感上升，个体对警察的满意度也会上升。在非西方国家，一些相似的发现也被证明(Alda, Bennett & Morabito, 2017)。

此外，社区的一些结构指标，建成环境特征和社区经济条件等与安全感的关系，也有不少学者进行了研究。Foster et al. (2014) 发现了土地利用混合性、街道链接性、公共巴士数量等建成环境特征与居民安全感的显著相关关系，土地利用多样、街道链接性好、巴士数量多的社区，居民拥有更高的安全感。而张延吉，秦波，和唐杰 (2017)，以及张延吉，秦波 & 朱春武(2019) 通过不同半径的建成环境 POI 指标得出土地利用混合型强、街道链接性好的社区，居民的安全感较低。也有研究认为，住宅类型与居民的安全感知有直接关系(Rollwagen, 2016)。

此外，一些研究也发现社区植被等因素与安全感知有关。一方面，由于封闭空间环境，植被被认为是恐惧感的触发因素 (Jorgensen, Hitchmough & Calvert, 2002; Herzog & Bryce, 2007; Maruthaveeran, 2016)。其中，一些研究使用图片来探测植被和安全感知的关系(Kuo, Bacaicoa & Sullivan, 1998; Baran et al., 2018)。也有一些使用实际问卷调查来探讨二者关系。Mak & Jim (2018)报告出，当访问在香港的公寓的时候，人们害怕会成为受害者。另一方面，一系列研究已经发现，人们的安全感知取决于植被的可视率。有好的视野可达性的树木被认为是与提升安全感有关(Kuo, Bacaicoa & Sullivan, 1998; Herzog & Bryce, 2007)。以上这些安全感大多是情景的安全感，这种安全感和稳定的安全感有区别。对于稳定的安全感，Maas 和他的同事证明了绿地数目和感知的邻里安全感之间的显著关系，以荷兰为例。即，生活环境中的绿地越多，人们的社会安全感就越高(Maas et al., 2009)。Mouratidis (2019) 以挪威奥斯陆都会区为例，证明了城市数目覆盖与增加的社区安全感的关系。

### 1.5 安全感知的测量与模拟

#### 1.5.1 安全感知的测量

纵观国内外文献，关于安全感知测量的分析较少。主要的相关研究都是基于调查问卷进行，因此本研究主要进行基于问卷调查形式的安全感知测评相关评述，再对近年来不断发展的其他测评方式进行介绍。

许多研究都通过询问受访者相关题项来测量安全感知，大多通过如下问题展开：“天黑后，你独自走在街上的害怕程度” (Ferraro & Grange, 1987; Hale, 1996; Doran & Burgess, 2012; Hodgkinson et al., 2017)，回答的选项主要包括“非常害怕，比较害怕，不太害怕，一点儿也不害怕”。然后，研究者将调查的样本推论到总体人口，分析不同群体的被害差异特征 (Hough, 1995)。

如今，学者们认为这类型的测评问题，安全感知就等同于基于主观概率的感知风险程度。这种测量方式实际上反映了受访者对待风险的态度，而不是一种实际经历的反映(Jackson, 2015)。这种方式测量的是一种静态的犯罪恐惧风险感知，或者是对未来遭受犯罪的焦虑程度 (Gray, Jackson & Farrall, 2011; Jackson, 2015)。

而安全感知常常是与日常生活中的个体经历息息相关的，具有复杂性特征 (Gray, Jackson & Farrall, 2008; Jackson & Gouseti, 2012)。

如果将安全感知作为在特定情景下经历的一种事件，这种作为事件的安全感知考虑到了个体日常生活的经历、特定环境、以及之前安全感等因素的影响。基于此，新的测评问题加入了频率和强度等维度 (Farrall et al., 1997; Farrall & Gadd, 2004; Gray, Jackson & Farrall, 2008 & 2011)。这些测量的问题集中于担忧情况，表示一直担忧的事件 (Farrall & Gadd, 2004)。这种转向表明，安全感知的测量是作为日常生活经历的事件，而不仅仅是一种潜在的焦虑的状态。

除了以上的风险和担忧维度，学者们提出安全感也包括行为维度。行为维度可以划分为避免性行为(比如，避免危险场所和人，在黑暗中在同伴陪伴下同行)和动员性行为(比如，安装户外的路灯，门锁，窗栏等)。类似地，Liska, Sanchirico & Reed (1988) 将这些行为区分为如下两类。一类是避免遭遇犯罪有关的情景和地方的行为，另一类是保护自己远离这些情景和地方的行为。

Gabriel & Greve (2003) 对以上的定义进行了补充。他们认为不安全感的状态的出现需要以下三个条件：个体关于被害的认知感知（认知维度），一个相应的情感经历（情感维度），以及一个适当的动机和行为倾向。这个观点也得到目前许多文献的认同，认为安全感(被害恐惧感)包括三个维度，分别是情感维度、认知维度和行为维度。关于不同维度对应的不同问题，如下表所示。

表 1-1 安全感知测量题项

维度	问题	具体描述	回答选项
情感维度	您自己担忧如下事件的频率如何？	我的家可能被盗；	从没有；
		我的财物有可能被偷；	偶尔；
		我会被人身攻击；	有时；
		我会被性侵或者被强奸；	经常；
		我可能在一次暴力事件中被杀；	特别频繁
认知维度	您认为这些事情可能在接下来 12 个月发生吗？	我的家可能被盗；	非常不可能；
		我的财物有可能被偷；	不太可能；
		我会被人身攻击；	很少可能；
		我会被性侵或者被强奸；	有可能；
		我可能在一次暴力事件中被杀；	非常有可能；
行为维度	在过去的 12 个月内，你的家里是否安置了任一如下事物？	防盗铃、门插销、额外的门锁、	是；
		防盗窗、保卫狗、户外安全灯、	否
		电子定时器、枪支	

参考 May, Rader, & Goodrum (2010) 和 Hanslmaier (2013)。

以上关于安全感知的测度都是以问卷问题作为测度方式。而近年来，随着传感器设备的发展，不少新的测量方式用于测量安全感。比如，Castro-Toledo et al (2017) 在丹麦一个城市做了实验，将心率作为被害恐惧感的一个心理指标，研究了街道灯光变化对受害风险感知的影响。Castro-Toledo 等人认为，心率作为表征受害风险感知的一个指标，反映了安全感的情感维度。而 Gabriel & Greve (2003) 认为受害风险感知行为维度的表现不仅包括计划的和控制的行为，还包括非计划的可见行为和生理反应。按照这个标准，心率则是作为受害风险感知的行为维度。此外，Guedes & Cardoso (2014) 利用眼动技术分析了影响安全感知的环境因素，发现照片中的表征社区失序的物体是影响被试安全感知的重要因素。而 Solymosi, Bowers, & Fujiyama (2015) 和 Solymosi et al. (2020) 利用手机 APP 应用回答相关问题测量安全感，这种安全感具有时间、空间特征，将安全感看做是一种事件。

然而，各种新型安全感知度量方式中，安全感知的具体定义不尽相同。一些测量手段实际上是测量了安全感的某个维度，而另一些测量手段实际上是测量的心理安全感，而不是犯罪学意义上的治安安全感。Jackson (2004) 区分了经验的不安全感和表达的不安全感，认为经验的不安全感是短暂的、频率低的。当居民认为犯罪威胁是可能的时候，经验的不安全感则出现。而表达的不安全感是人们对社区和更广泛的社会文化的感知而构建的。

综合来看，基于不同的测度工具和维度，居民安全感知的测度方式主要如下表所示。

表 1-2 安全感知的不同测度工具

测量工具	常用测量维度	类别	代表文献
调查问卷	情感、认知、行为维度	治安安全感	Hanslmaier (2013)
手机 APP 应用	情感维度	治安安全感	Solymosi et al. (2015)
医学设备	情感或行为的生理反应	治安安全感 心理安全感	Castro-Toledo et al. (2017)
Twitter 文本	情感或行为的言语反应	治安安全感 心理安全感	Williams et al. (2017)
图片	空间环境的心理感知	心理安全感	Guedes et al. (2014)

### 1.5.2 基于时空视角的安全感知测量

日常活动理论把焦点从犯罪者动机转向到犯罪者事件 (Cohen & Felson, 1979)。这个理论第一次同时考虑了犯罪的时间和空间尺度。类似，安全感也从状态转向到事件，把安全感作为一种事件，可以寻找到与这些事件相关的环境，从而利用情景犯罪预防方法减少不安全感。我们都在生活历程中，要走出恐惧的阴影，受到自身的经历以及空间、时间和社会环境的影响 (Pain, 2000)。可以假设，即使一个短的时间段，人们会经历不同水平的安全感知。因此，测量这种变化，需要时刻跟踪个人，因为个体每天经历不同的时间、空间、社会环境。此外，安全感的测量还应该关注有关情景因素的调查。犯罪地理学强调，犯罪活动中的决策过程是被个体和环境因素影响。与此类似，一个被害恐惧事件也受到个体和环境因素影响，学者们强调个体的安全感知受他们所处的社会和建成环境特征的影响 (Wyant, 2008; Lis, Pardela & Iwankowski, 2019)。

与警务犯罪数据不一样，不安全感事件经常没有特定时间和地址，因此安全感的空间化通常是聚集到一个特定社区，比如邻里。Solymosi, Bowers, & Fujiyama (2015)利用伦敦市 borough 空间尺度的 MPSPAS 调研数据对居民安全感进行了空间化。MPSPAS 数据中每个辖区的调研样本为 268-681 个，但是很多 boroughs 有超过 300000 人口。这表明，如果调研人口不多，是否能够真实代表这个区域的人口是有风险的。事实上，如果样本量足够，居民安全感的空间测量还可以在更小的尺度上展开，比如住宅小区。很多研究探讨了社区失序与居民安全感的关系，发现社区的失序程度是居民治安安全感的重要因素 (Franklin et al., 2008)。因此，测量邻里的失序水平，也会反映出邻里居民治安安全感的高低程度。但是，社区失序在社区内也是不一样的。Weisburd (2012) 指出，高社区失序的街道通常被低社区失序的街道环绕。因此，这表明不安全感事件并不是完全均匀分散在邻里，而是在空间上集中在微观的地理区域，比如街道或者角落处。借助移动手机设备，一些研究通过被试者连续一段时间的实验观察，发现基于微观时空尺度的安全感特征 (Solymosi, Bowers, & Fujiyama, 2015; Chataway et al., 2017; Solymosi et al., 2020)。

因此，对于安全感来说，不同地理尺度的测量方式不同。个体水平上，安全感相当大程度上受个体经验的影响，而邻里水平上的安全感主要受人们生活环境

的影响。在更大的水平上，安全感知是一种社会现象，一种一般化的焦虑状态。因此，安全感知测量的尺度非常重要。

这些基于社区的安全感知研究通常是研究居住区的安全感知，而人们通常会有超出居住区的日常活动，比如工作地，大学校园，公共交通场所。因此，一些研究者也使用问卷来调查特定地区环境与安全感知的关系，以及开展参与式制图和定性 GIS 的研究。利用问卷调查，让被试者在地图上画出感觉不安全的地点，Pánek, Paszto, & Marek (2017)对捷克 Olomouc 市的安全感知状态进行了制图，发现了城市一些安全感低的区域，并进行了一系列空间分析。

### 1.5.3 安全感知的空间模拟

地图学研究往往关注可视化或者能够被可视化的要素，比如温度、风速等。批判主义和女性主义地理学者倡导制图人们感受的场所，其中包括情感。安全感知制图是地理信息科学、环境科学、城市规划、心理学等领域的前沿研究话题。由于安全感的指标多来自于实地调研收集，因此根据已有的调研数据模拟出更大范围的安全感空间分布状况尤为必要。

伦敦大学学院的 Williams et al (2015 & 2016) 利用不同时期的基于行政单元的公众对警务的满意度数据，使用时空贝叶斯模型，预测了警务满意度的时空分布状态，并取得了较好的预测精度。利用 1300 个美国郡县的 Twitter 的推文数据，Schwartz et al. (2013) 探索了能否通过这些推文数据预测基于问卷调查得到的主观安全感空间分布特征。研究发现，使用 LDA 算法提取的词频可以提高生活满意度预测的准确性。户外活动、精神愉悦、锻炼、良好的工作等词汇与生活满意度的提高有关，而表示无聊或者疲倦的词汇与生活满意度呈负相关。此外，麻省理工学院 MIT 的一系列研究，通过公众参与的方式，让用户在线对 Google 街景图像的不同维度打分（包括道路安全感、道路清洁度、美丽度等），最后获得了约十万张照片的不同维度的主观环境感知评分，然后通过深度学习方法对海量照片进行自动评分，绘制了 Linz, Salzburg, Boston, NYC 等一系列欧美城市或地区的城市感知地图 (Salesses, Schechtner, & Hidalgo, 2013; Naik et al., 2015)。Zhang et al.(2018) 也运用类似研究方法，对北京和上海等地的城市感知地图进行了精细化刻画，并据此深入分析了环境和感知的关系。利用街景数据源，国内学者近

年来进行了一系列工作,分析城市空间品质(唐婧娴& 龙瀛,2017; 叶宇等,2019; 崔成等, 2020)。

综合来看,运用社交媒体数据或者开源数据来进行的情感制图研究,近年来发展迅速。但是对于基于问卷调查的治安安全感的空间模拟研究,特别是基于社交媒体数据的治安安全感空间模拟研究,还没有研究涉及。这需要根据特定的空间尺度,科学探索如何实现基于开放数据源对传统调研获得的个体治安安全感状况进行空间模拟。

### 1.6 国内外研究述评

安全感知研究一直是地理学、犯罪学、环境科学、计算机科学等多学科交叉的热门话题。整体上,关于安全感知及其关键要素的测量,学者们开展了大量有价值的研究工作,并且取得了具有理论意义和现实意义的研究成果。

国外来看,大量研究集中在个体和环境因素如何影响安全感知,安全感知的空间模拟也在不断发展。国内来看,安全感知的文献还相对缺乏,除了不断有相关学者探讨安全感知的影响因素(林冬阳 & 田宝江, 2014; 彭玉伟, 2015; 肖鸿元, 2017; 王珂, 2017; 朱志玲, 2020),客观自然和建成环境与安全感知关系的研究还很少,同时还缺乏研究开展安全感知的空间模拟。

通过对安全感知相关环境要素提取、影响因素分析、安全感知空间模拟等方面研究的文献梳理,发现当前的研究还存在如下一些问题。

#### (1) 安全感知相关环境要素提取: 传统环境要素度量方法有待突破

城市街景环境特征的精细化识别,对进一步研究环境和安全感知关系有重要意义。安全感知相关环境因素的测量主要是基于传统的问卷手段,传统手段存在时效性低、成本高等固有缺陷。随着开放数据的丰富,不断有利用政务大数据、网站大数据、文本大数据、街景大数据等多源数据开展建成环境或社会环境测量的新方法。深度学习技术在计算机领域发展迅速,基于机器学习技术的图像识别、图像分类已经有成熟的研究模型和范式。由于街景数据在信息丰富性、科学性方面的优势,有较多的研究尝试利用街景数据识别城市环境要素。但是,当前基于街景影像的环境要素识别主要集中在方法的改进上,针对特定城市的大规模街景识别应用还比较少,特别是在环境有序/失序领域的应用研究还较少。因此,借助

于深度学习技术，准确地对特定大城市街景环境/失序特征识别，建立基于街景影像的失序微环境要素评价体系，是一个有意义的实际问题。

### **(2) 环境和安全感知的关系：安全感知的环境影响因素还需要进一步分析**

环境对安全感知的影响一直是一个热点的研究话题。虽然以往研究揭示出一些环境特征，如人口密度、植被、住宅环境等因素可以影响居民安全感知。但是，由于技术限制，一些环境因素（如绿视率、区域商业设施数目、区域天空开阔程度等）与安全感知的关系仍然没有被探讨。但是，这些因素不能被轻易忽视。依据 CPTED 环境设计预防犯罪理论，天空开敞空间小的区域、绿视率低的区域可能和更高的安全感知有关。因此，探索并验证更多微环境因素与安全感知的关系，不仅可以丰富安全感知的理论研究，也可以对提升居民安全感的实际工作有有效指引。

### **(3) 基于街景图像的社区环境指标对于安全感的影响机制还有待探讨**

根据环境设计预防犯罪理论，虽然有研究表明一些客观环境指标是影响居民安全感知的重要因素。但是，关于客观环境作用于安全感知的机制研究还很少涉及。少数研究（Brunton-Smith, Jackson & Sutherland, 2014）提出了客观环境通过感知的社区失序和社区融入作用于安全感知的机制。因此，在探讨基于街景的微环境与安全感知关系的基础上，研究客观环境-主观环境-主观安全感知的影响机制，有利于进一步分析出安全感知产生的内在机理。这是一个有理论意义的研究话题。

### **(4) 安全感知测评与模拟：研究方法有待更新**

目前绝大多数的安全感测评研究都还是通过问卷调查的手段获取，这种调查方式虽然具有准确性高、操作简单等优点，但是这种调查方式不适宜大规模展开，同时时效性慢，时间、金钱、成本大。目前，我国已经进入全面建成小康社会阶段，人民群众的生活品质提高尤为重要，各级政府部门都在广泛开展公众满意度、公众安全感调查，以保障社会稳定、国泰民安。因此，对于安全感知的大规模快速评价的需求尤为旺盛。目前，虽然有各类研究尝试通过手机 APP、心率、Twitter、街景影响等各种方式测量心理安全感，但是这些方式由于全面性、样本量、科学性等局限，都不能代替问卷安全感的测量。而在相关研究领域，一系列研究以开放数据源为预测变量，以基于问卷的测量值为结果变量，已经取得了显著的成果。

Jean et al. (2016)在 *Science* 期刊发文，利用夜间灯光影像和机器学习算法，模拟出了非洲大陆各国的消费、收入、资产程度，模拟结果与问卷调查的结果高度相关。这种方法为科学、客观、低成本地测度出非洲各国发展水平、促进非洲大陆经济发展具有重要意义。Gebru et al. (2017)在 *PNAS* 美国科学院院刊发文，通过大规模的街景数据和深度学习算法，提取出图片中不同型号的汽车，通过统计分析发现基于街景数据的汽车指标可以有效反映出美国各郡县的社会经济情况。因此，在安全感知研究领域，借助多源大数据，如何有效、准确地刻画基于问卷的安全感尤为重要。这样，以后的社区安全感调研，通过小范围抽样社区的调研即可科学客观模拟出大规模社区的治安安全感状况，有利于政府提升城市社会治理的效率。

## 第二章 研究问题与研究设计

针对前两章节的铺垫，本章节提出研究问题，并开展一系列研究设计。本章节是本论文的框架性内容，承上启下，为进一步的实证研究设计研究路线和研究流程。

### 2.1 研究问题

本研究以居民的个体安全感知为研究对象，从社会生态学的视角出发，研究安全感知模型的影响因素及其空间模拟。本研究的研究问题主要有三方面（表 2-1）。首先，本研究尝试利用街景数据挖掘出一系列微环境因素，尝试将这些微环境因素加入安全感知模型，检验这些微环境因素与安全感知的关系，以及检测微环境因素的加入是否会提高安全感知模型的拟合效果。其次，本研究分析相关微环境因素对安全感知的内在影响机制，分析相关微环境因素是否不仅有直接作用，还有间接作用。下一步，如果相关微环境要素对安全感知既有直接作用，也有间接作用，那么如何科学选取预测模型进行个体安全感知预测？最后，本研究探讨是否可以开发一套从小样本问卷调查到大规模空间模拟的安全感知空间模拟方法。具体来看，

(1) 街景微环境与安全感知的关系：街景微环境的加入是否会提高安全感知预测模型的解释能力？

根据日常活动理论，基于路网的街景微环境指标能够更加真实反映社区居民的日常环境暴露，进而影响居民的社区感知。相较于传统指标，街道微环境指标更有可能影响安全感知。然而，传统手段很难精细刻画城市的微环境，虽然不少研究分析了主客观环境与居民安全感知的关系，但是还没有研究探讨街景微环境因素与安全感知的关系。因此，本研究利用开放数据和深度学习算法，拟获取广州市域范围内的街景微环境特征，试图分析街景微环境的加入是否可以提升安全感知模型的解释能力，以此分析街景微环境与安全感知的关系。

(2) 街景微环境对安全感知是否不仅有直接作用，还有间接作用？

虽然目前已有一系列研究探讨了环境（包括社会环境、自然环境、感知环境）对安全感知的影响。但是，很少研究建立连接“客观环境-主观环境-安全感知”的

影响机制分析。在研究相关街景微环境因素与安全感知关系的基础上，本研究试图进一步分析主客观环境如何相互作用，从而导致了个体安全感知的变化。同时，这个研究问题也回答了街景微环境因素对安全感知是否不仅有直接作用，还有间接作用。

(3) 传统的安全感知模型都是线性模型，是否非线性安全感知模型具有更好的预测能力？

当前研究中，绝大部分安全感知模型是线性模型开展的。然而，安全感知作为一种感知，实际上感知的变化并不是严格线性变化的。在一定阈值区间外，安全感的变化可能是非线性的。而且，安全感知的影响因素复杂，其中一些因素可能同时存在直接作用和间接作用，一些因素可能只存在直接作用或间接作用，这种情况下使用非线性的模型可能有利于安全感知的预测工作。因此，本研究将比较非线性模型和线性模型对安全感知的预测效果，探索出一种更准确的安全感知预测模型，为实际的安全感知预测和空间模拟准备。

(4) 如何开展安全感知的空间化模拟？

传统的居民安全感调查都是通过问卷调查的形式开展。由于时间、精力、成本巨大，这种方式不适合大规模开展。本研究拟开发一套安全感知的空间化模拟方法，尝试如何先利用问卷数据建立个体安全感预测模型，最后脱离问卷数据实现研究区所有个体的安全感知预测，进而实现安全感知空间模拟。

关于以上几个研究问题的内在联系，如图 2-1 所示。其中，首先计算相关街景微环境指标，以作为整个研究核心的数据支撑。研究问题 1 和研究问题 2 分别探讨了街景微环境因素与安全感知关系之间的直接效应和间接效应。问题 3 是问题 1 和问题 2 的深化，对比了线性方法和非线性方法对于安全感知预测的效果。问题 4 是问题 3 的实际应用示范，也是本研究的最终问题。

表 2-1 研究问题总结

内容	研究现状	研究不足	研究问题
相关环境要素测量	①客观环境主要基于遥感影像或测绘技术获取; ②社会环境和环境失序主要基于问卷或统计手段获取; ③近年研究利用街景大规模评价街区失序、主客观环境评分	①没有研究大规模提取并分析广州市域范围的多维度环境有序/失序特征	①如何刻画广州市域范围的多维度环境有序/失序特征
安全感知因素与机制	①相关研究多是基于传统的环境指标研究环境与安全感知的关系	①没有研究考虑基于街景的微环境指标(可以更加真实地刻画个体的日常环境暴露)对安全感知的作用; ②很少研究分析客观环境-主观环境-安全感知的机制	①微环境指标是否对安全感知有直接作用? ②微环境指标是否对安全感知有间接作用?
安全感知空间模拟	①主要是基于调查问卷的空间制图; ②有新方式尝试,如利用手机设备等进行时空安全感测量; ③基于犯罪率进行安全感知感建模;	①没有研究开展大范围的安全感知空间模拟	①如何基于少批量的问卷调查,结合多源数据,模拟出城市社区安全感知的空间分布?

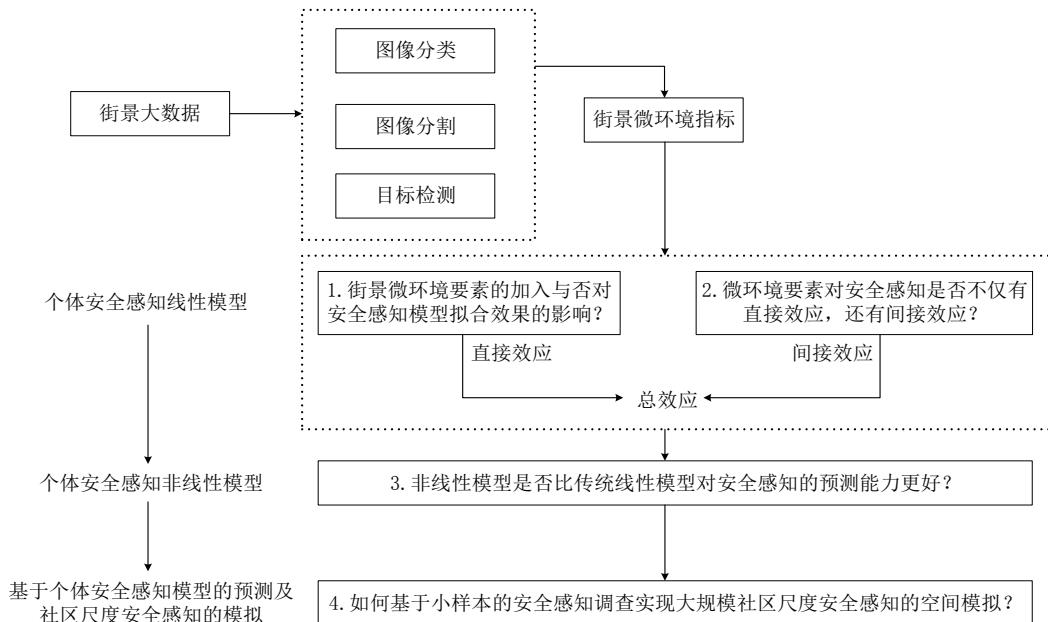


图 2-1 研究问题的内在逻辑

## 2.2 研究区域

本文的研究区域为中国南方的广东省广州市,坐落于粤港澳大湾区核心区域,

毗邻南海。广州是中国第三大城市，城市化率为 85.53%，2019 年的常住人口超过 1500 万<sup>1</sup>。广州市包含 11 个区级行政区域，面积大约为 7434km<sup>2</sup>。其中，根据目前的区划，7 个中心城区包括越秀、海珠、荔湾、天河、黄埔、白云、番禺，其余区域是外围区。

地貌上，广州位于  $112^{\circ}57'E - 114^{\circ}30'E$ ,  $22^{\circ}26'N - 23^{\circ}56'N$  之间，北回归线从广州市北部穿过。地形从东北向西南降低，气候上属于亚热带季风气候，气候温暖、降雨充沛。年均温度在  $20-22^{\circ}C$ ，广州市区年均降雨为 1720 mm。水文方面，广州市区划范围内河网密布，水系发达，由北江、东江、西江组成的珠江水系纵贯广州。

社会经济上看，广州市是我国的国家中心城市之一，也是 4 大一线城市之一，2020 年广州市的 GDP 为 25019.11 亿元。广东省是中国人口流入最多的省份，作为广东省的省会，广州市是中国制造业的重要中心。从 2015 年开始，广州市的流动人口已经超过城市总人口的一半。这个比例更是逐年上升。

选择广州市作为本文的研究区有以下几点考虑。首先，广州市的经济基础较好，流动人口较多，社会异质性强，可以充分检验国外安全感知的理论，开展安全感知的研究。其次，广州市的气候温暖，全年温差小，环境因素变化相对较小，有利于检验环境和安全感知的关系。最后，广州市政务公开工作开展深入，各项基础数据资源完善丰富，有利于研究的各类数据获取。

在开展具体研究时，广州市的范围有所缩减。因为街景和手机数据等开放数据在广州中心城区覆盖全面，而在增城、从化等郊区的覆盖度不完善。所以，在研究开展阶段，本研究的研究区域不包括花都、南沙、增城、从化等四个区域。

### 2.3 研究内容及总体技术路线

#### 2.3.1 研究内容

总体来看，本研究的研究内容有五点：

(1) 基于街景的安全感知环境影响要素提取及制图

主客观环境对于安全感知具有重要的影响作用。利用开放数据源，本研究基

---

<sup>1</sup> [http://tjj.gz.gov.cn/gkmlpt/content/5/5727/post\\_5727607.html#226](http://tjj.gz.gov.cn/gkmlpt/content/5/5727/post_5727607.html#226)

于编程获取大批量街景影像，综合运用多种深度学习算法获取街景环境指标，进而进行空间格局分析，识别出广州市失序环境要素的空间结构特征及其差异。

### (2) 街景微环境因素对安全感知的直接作用分析

针对基于日常活动理论的多类型环境因素，本研究运用多元回归统计技术探测这些微环境暴露因素对于居民安全感知的影响机制。本研究比较微环境要素的加入与否，对安全感知模型拟合效果的影响。同时，本研究比较多种社区环境失序的度量指标，分析客观环境失序对安全感知的作用。

### (3) 环境因素对安全感知的影响机制分析

环境因素对于安全感知的影响是一个复杂的机制。为了更清楚地挖掘这个机制，指导进一步的安全感知预测，本研究设计了从客观微环境到主观环境再到安全感知的研究思路。并且探讨感知环境因素在客观微环境对安全感知的影响机制中的中介效应、调节效应，以及有中介的调节效应。总体来说，在街景微环境要素的直接效应基础上，本研究同时探讨街景微环境要素的间接效应。

### (4) 个体安全感知的预测

传统的安全感知分析基本是基于传统线性回归模型开展。在探测出街景微环境因素的直接作用和间接作用基础上，本研究综合运用多种非线性方法（机器学习算法）和线性方法，比较非线性模型和线性模型对安全感知模型的预测效果。

### (5) 安全感知的空间模拟

传统的安全感知的空间化是基于问卷开展。通过不同社区等区域的一定样本抽样调查居民的安全感知，然后聚合个体安全感知到空间安全感知。本研究，在多源实时数据的驱动下，综合比较多种机器学习算法，选择最优模型建立了个体安全感知的预测模型。最后，运用科学模拟的广州市各社区居民的属性数据，分析研究区所有个体的安全感知，进而完成安全感知的空间模拟和制图分析。

### 2.3.2 总体技术路线

本研究的总体研究路线为，利用基于街景数据的多维度微环境指标，检验微环境与安全感知之间的关系和机制。根据相关影响因素，建立安全感知预测模型，并完成安全感知的预测，实现安全感知的空间化。详细的研究内容如图 2-2 所示。

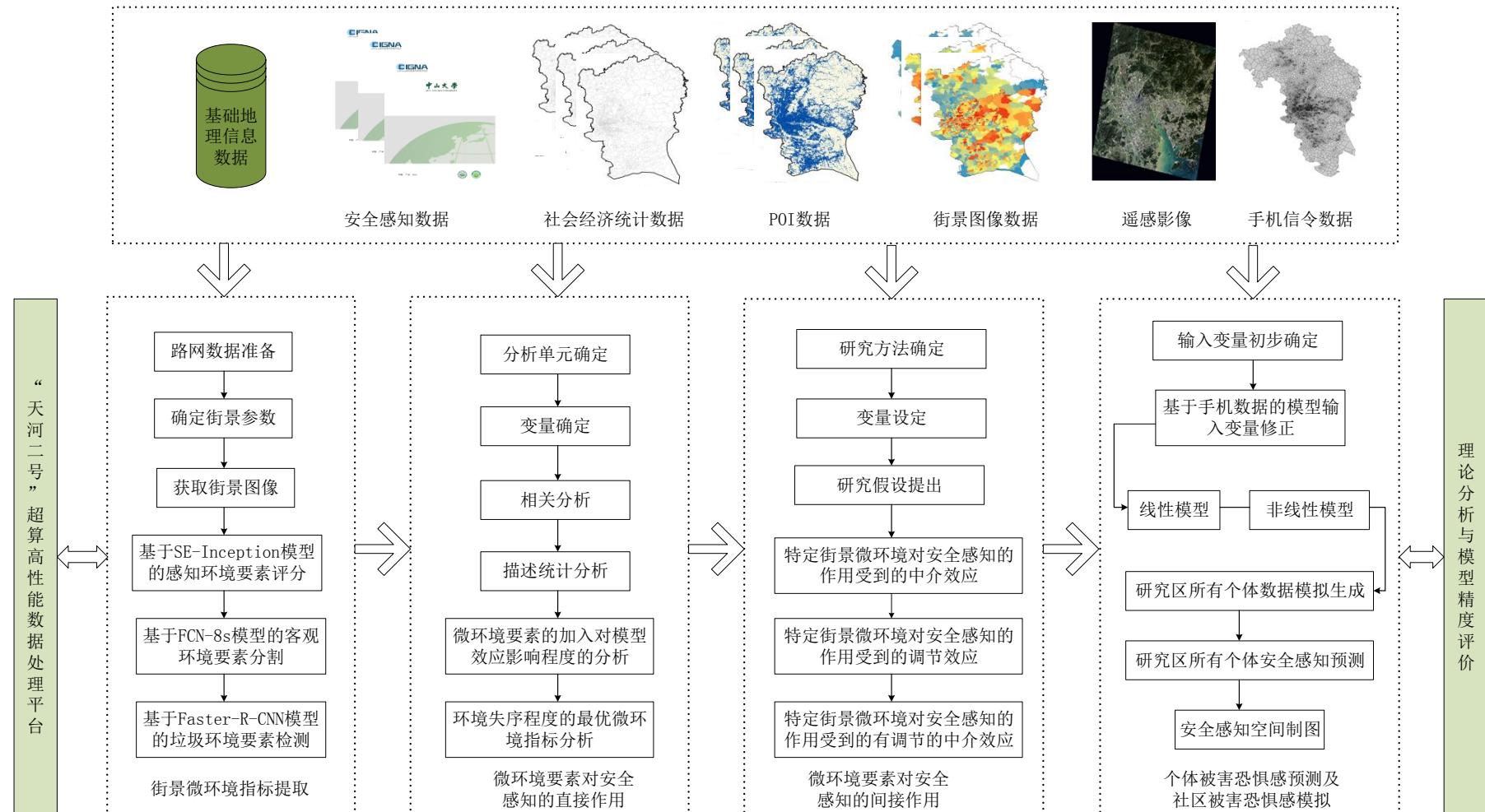


图 2-2 技术路线图

### 2.3.3 论文框架

本文的论文框架详见图 2-3，研究一共分为八章，逻辑上是从背景分析→研究设计→关键要素制图→影响因素分析→影响机制分析→空间化模拟，由浅到深，逐步探讨安全感知及其空间化模拟的话题。

第一章，绪论。本章包括研究背景、意义和概念界定等内容。

第二章，文献综述。本章包括综述前人在安全感知空间分布及模拟、环境要素测量、安全感影响因素及测量等方面的研究基础，进而提出研究不足。

第三章，研究问题和研究设计。本章主要提出了研究问题和研究设计方案，并且对主要的研究数据进行说明。本章节内容是本研究的框架。

第四章，基于街景图像的环境因素提取。本章分别采用基于卷积神经网络模型 SE-Inception、全卷积神经网络模型 FCN、目标检测模型 Faster-R-CNN 识别了街景图片的多维微环境要素。运用空间自相关分析方法，进一步探讨广州微环境要素的空间分布特征和空间集聚特征。

第五章，基于上一章节获取的环境要素数据，本章节运用统计学中的最小二乘法多元线性回归模型，在进行一系列数据预处理的基础上，考察了基于安全感知的环境要素分析。本章节主要分析微环境要素是否是解释安全感知的重要因素，以及哪种微环境指标是度量社区环境失序的最佳指标。

第六章，环境因素对安全感知的影响机制研究。该章节综合运用中介分析、调节分析、交互分析、有调节的中介分析等多种分析手段，研究客观环境-主观环境-安全感知之间的影响机制。本章节主要分析微环境要素对安全感知的间接作用。

第七章，在之前章节的基础之上，该章节的目的是对安全感知进行预测，然后进行空间化模拟。本章节拟用随机森林算法、SVM 支持向量机算法、二元 Logistics 回归模型算法，比较得出最优的算法。最后，根据统计年鉴数据模拟出研究区所有人口的安全感知相关属性数据，并对研究区范围内的个体进行个体安全感知模拟以及空间化制图，以此指导具体的城市管理实践。

第八章，该章节主要包括总结与展望，概述整个论文的结论、创新和研究不足。

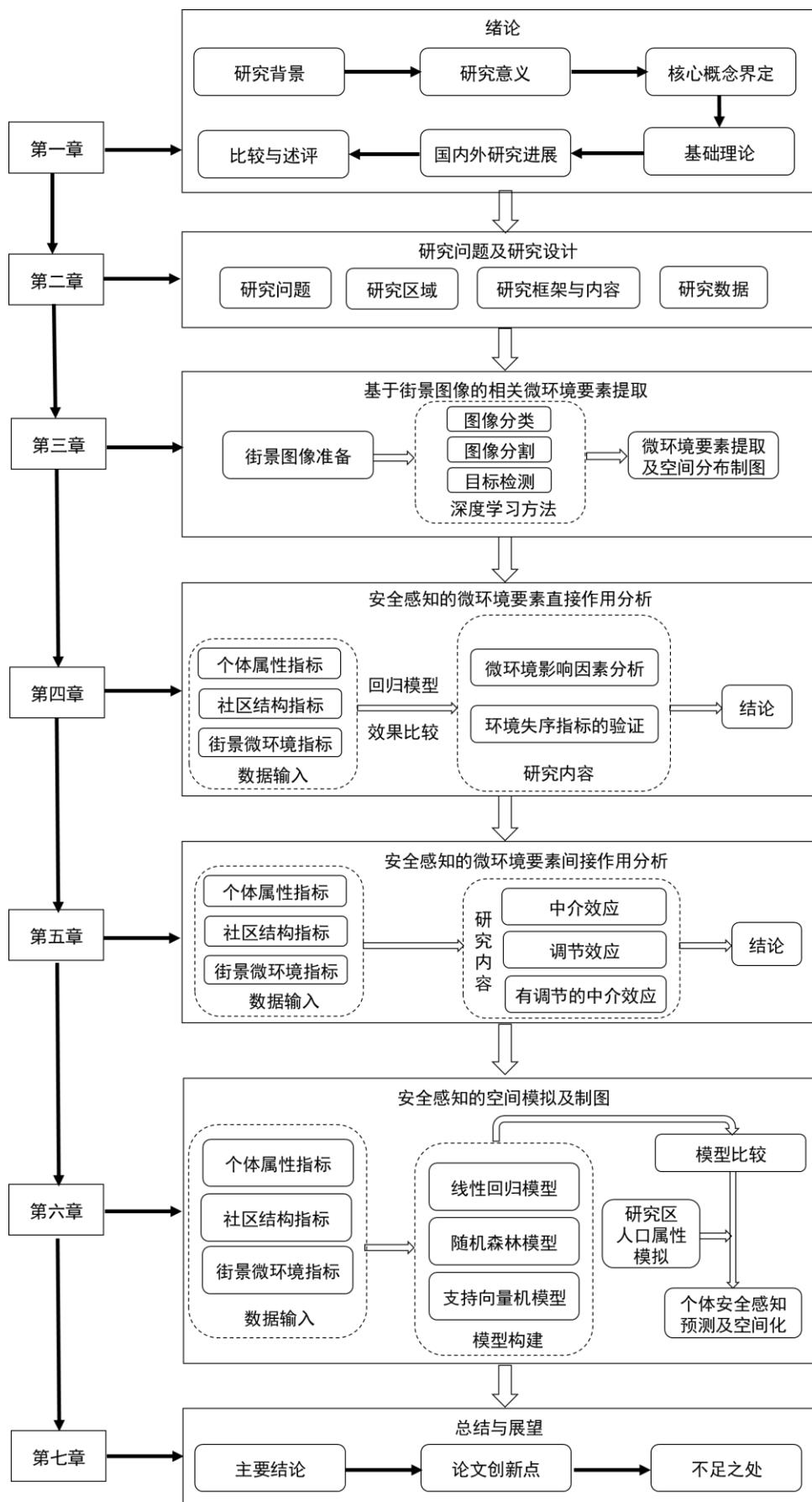


图 2-3 论文框架流程图

## 2.4 研究数据

### 2.4.1 广州市路网数据

本文采用广州市路网数据（图 2-4），来源于购买的道道通公司的广州市路网数据。道道通公司为百度地图提供导航电子地图及其相关技术产品。因此，相较于 OSM 等开放地图数据，本研究中的道路路网数据更加丰富、准确。路网数据包括辅路、高架、高速、国道、环城高速、其他道路、省道、市区二级、市区一级、市区杂路、县道等各种类型的道路。

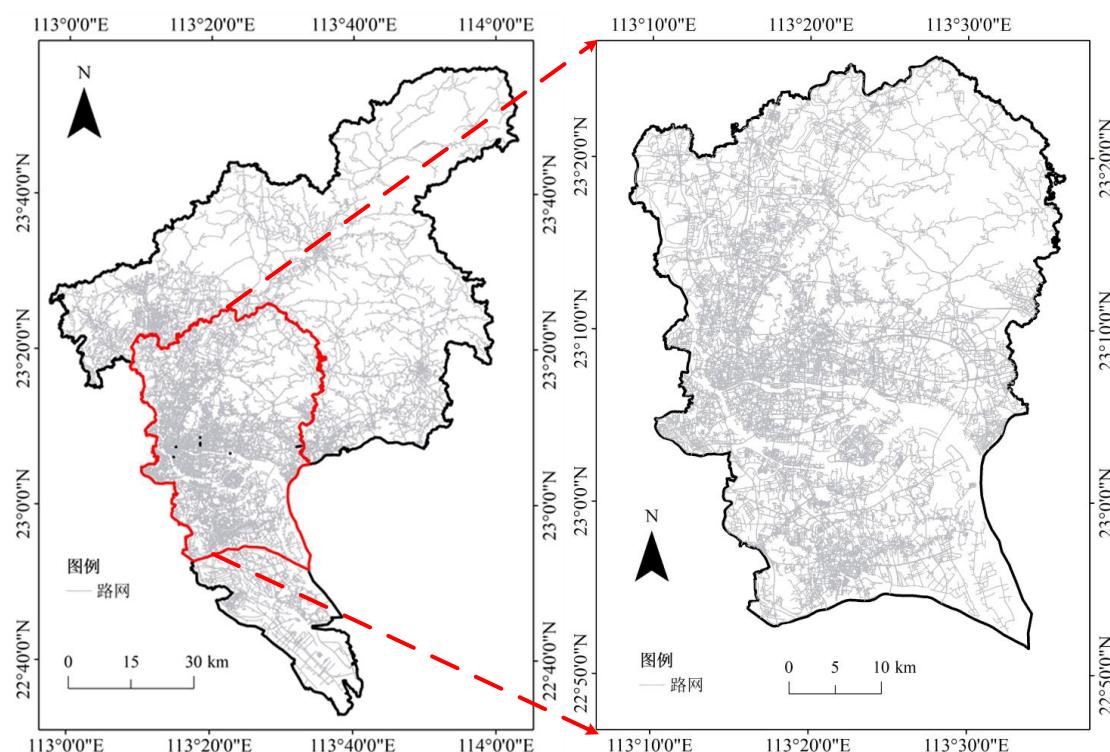


图 2-4 研究区路网地图

### 2.4.2 百度街景图像数据

本研究从百度地图开放平台获取了 4 个角度约 1023020 张街景照片，作为街景指标计算的源数据。数据的采集时间为 2018 年 12 月 1 日，采集时间耗时约 1 个月。

#### 2.4.2.1 百度街景参数介绍

街景车配备有摄像仪器、激光仪器和 GPS 定位仪器，当街景车以一定速度开

过时，可以获取周围的全景图像。通过后期处理和生产，能够展示出全景图画面。百度地图是中国领先的地图服务提供商，其提供的街景地图服务覆盖中国绝大多数城市。根据特定位置的坐标，该位置的街景图像可以通过百度地图提供的全景静态图 API 服务以 HTTPURL 的形式调用。在百度地图开发平台的示例中（<http://lbsyun.baidu.com/index.php?title=viewstatic>），如果要请求奎科科技大厦周边点（地理坐标为东经 116.313393,北纬 40.04778）的全景图，则在浏览器上输入一个 HTTPURL 地址：

[http://api.map.baidu.com/panorama/v2?ak=E4805d16520de693a3fe707cdc962045& width=512&height=256&location=116.313393,40.04778&fov=180](http://api.map.baidu.com/panorama/v2?ak=E4805d16520de693a3fe707cdc962045&width=512&height=256&location=116.313393,40.04778&fov=180)

然后，浏览器会自动返回该参数下的街景图片。



图 2-5 该地址街景图片

该 HTTPURL 地址包含一系列参数设置。参数包括：

location（采样点坐标。它的输入参数一般是经纬度坐标值，百度地图 API 会自动根据坐标位置捕捉半径 50 米内最接近该位置拍摄的全景图。）

size（该参数确定图像的输出大小，以宽度×高度的形式输出图像大小。）

heading（该参数表示摄像机的方位，值区间为[0, 360]。默认值为 0。0 和 360 两个值都表示正北方向）

FOV（水平方向范围。该参数表示图片视场角，系统值为 90。视野单位以度为单位表示，取值区间在[0,360]，fov=360 时可以显示整幅全景图。）

pitch（该参数表示摄像仪器相对于街景车的俯仰角度，取值区间为[0,90]。默认值为 0，即平视。正值表示摄像仪器向上倾斜，负值表示摄像仪器向下倾斜。）。

ak(用户的访问密匙。需要从百度地图开发平台申请，不同类型用户 ak 的权

限不同，涉及每天能够顺利返回的街景图片次数。)

coordtype(位置点的坐标类型。可以设置为百度坐标、GPS 坐标、以及其他类型坐标)。

#### 2.4.2.2 路网采样点布置

根据成本和运算时间考虑，本文设置 50m 为道路采样点的间隔，以保证完整记录道路两侧地物的变化。利用 ArcGIS10.3，完成了采样点的生成。广州市整个行政区划共有 416345 个采样点，本文研究区（广州市七个中心城区）共有 238028 个采样点，分布如图 2-6。

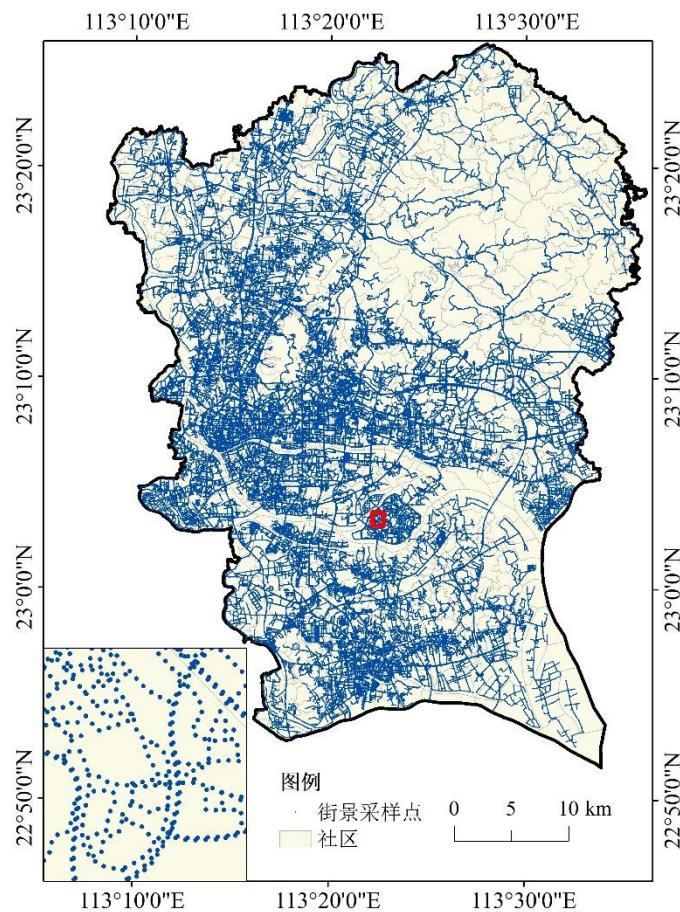


图 2-6 广州市研究区街景采样点

### 2.4.2.3 镜头设置和图像下载

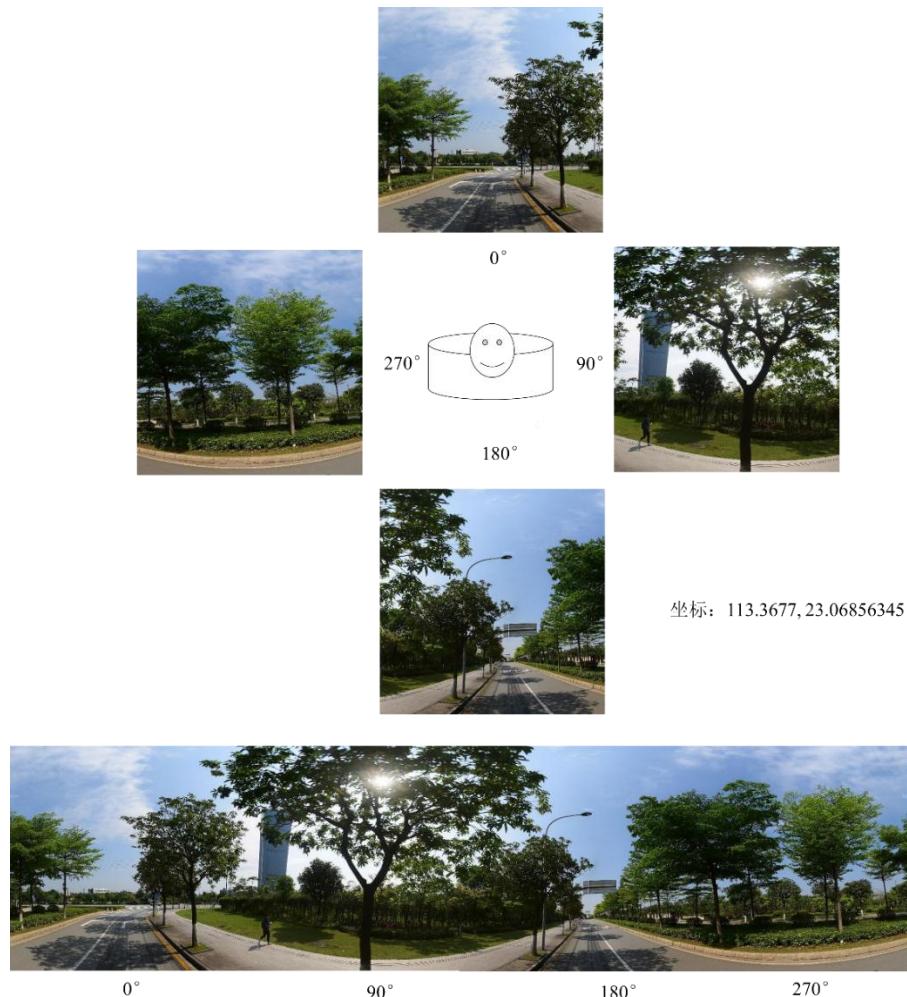


图 2-7 不同 heading 参数设置示意图

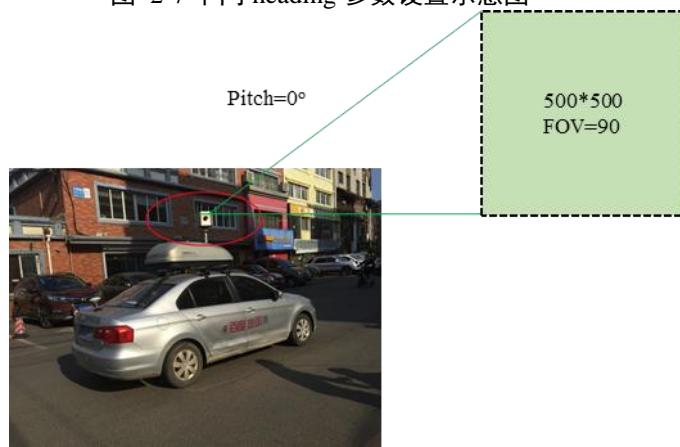


图 2-8 Pitch 参数设置示意图

为了能够下载得到全景的静态图像（图 2-7），本文需要对镜头和拍摄朝向进行设置。街道环境下，行人的视野范围为  $360^\circ$ ，故 heading 参数要覆盖四个角度。关于 Pitch 参数，由于本研究区域为广州市大部分范围，出于数据量大小考虑，

本研究只考虑最优情况下行人水平观察前方的角度，即 Pitch 参数设置为  $0^\circ$ （图 2-8）。本研究中 fov 值设置为默认值，即 90 度。

本研究基于 Python 编程语言，完成图片抓取。由于百度 API 每日有最大下载量限制，在 2018 年底个人开发者的最大下载限制是 10000 次调用。本研究申请了 5 个不同账户，共计花费约 1 个月完成照片的采集。采集时间为 2018 年 12 月。下载到广州市行政区划内 255758 个采样点的 4 个角度的照片，共计 1023032 张。研究区内有 195830 个采样点的 4 个角度的照片，共计 783320 张。

综合考虑采样精度、运算效率，本研究将图片输出的分辨率设置为  $500 \times 500$ 。将全部路网样点的坐标信息保存在 Excel 文件中，利用 Python 读取采样点的位置信息，获取 Location 参数，生成 HTTPURL 地址信息，再自动解析 HTTPURL 地址来下载该采样点的图片信息。详细镜头参数设置如下表所示。

表 2-2 不同街景镜头参数设置详细介绍

<b>Location</b>	<b>Size</b>	<b>FOV</b>	<b>Heading</b>	<b>Pitch</b>	<b>Coordtype</b>
			0		
采样点经纬度	500×500	90	90	0	WGS84II
			180		
			270		

### 2.4.3 问卷调查数据

本研究使用研究团队于 2016 年 1 月至 4 月研究设计，并通过全国领先的知名问卷调查公司“豪森威”<sup>2</sup>发放，耗资 20 余万获取的广州社区环境与居民安全感项目（Project on Public Safety in Guangzhou Neighborhoods, PPSGN）调查问卷（如图 2-9）。PPSGN 安全感知问卷包含多维度的问题，如个人基本情况、房屋基本情况、社区基本情况、社区安全及满意度、报案率等题项。

基于成本和样本代表性的考虑，问卷拟收集 2000 份。具体的调研方法如下。第一，使用主成分分析（PCA）从六普统计年鉴中的 30 多项数据提取出关键特征，然后使用聚类分析法识别社会区。这样，所有广州市社区被划分为 9 类社会区，包括本地老龄化社区，老城区中产阶级区域，大学城区域等。城市社会区域是指居住在某一区域内，生活水平、生活方式、生活状态大致相同的同质人口的聚集。生活在不同社会区的人们有不同的特征，信仰和行为。反映在空间上，社

<sup>2</sup> <http://www.hswell.com/>

社会区是由多个社区组成的城市同质区域。早在 1950 年左右，一些学者以旧金山为例开始研究社会区 (Shevky & Williams, 1949)。划分社会区的目标是进行分层抽样，使样本能够代表广州不同类型的社区。其次，根据各社会区人口占广州市总人口的比例确定各社会区的样本量。第三，采用等概率选择方法对不同的样本社区进行系统采样。在本次调查中，一共 90 个样本社区被选择。这样，虽然样本量不是很大，但是这些样本全面代表了广州市不同类型的人口。

PPSGN 问卷数据集的所有调查者都是受职业训练后进行调研，以保证问卷的质量。问卷呈现在平板电脑上，所有调研者辅助调查者以完成问卷调查。为了保证受访者的参与度，需要确认受访者是否匹配调查的要求（大学生以及年龄低于 18 岁的个体被礼貌谢绝参与问卷），因此只有每一户中合格的成年人被纳入问卷调查中。每个社区获取了按计划的抽样配额。样本的回复率为 87.4%，与国内的抽样调查样本回复率持平(Liu et al., 2009)，与国外相比偏高。在排除错误和有问题答案后的问卷，PPSGN 一共 1994 个有效的样本被收集，样本有效率为 95.1%。



图 2-9 社区安全感知问卷

#### 2.4.4 六普人口统计数据

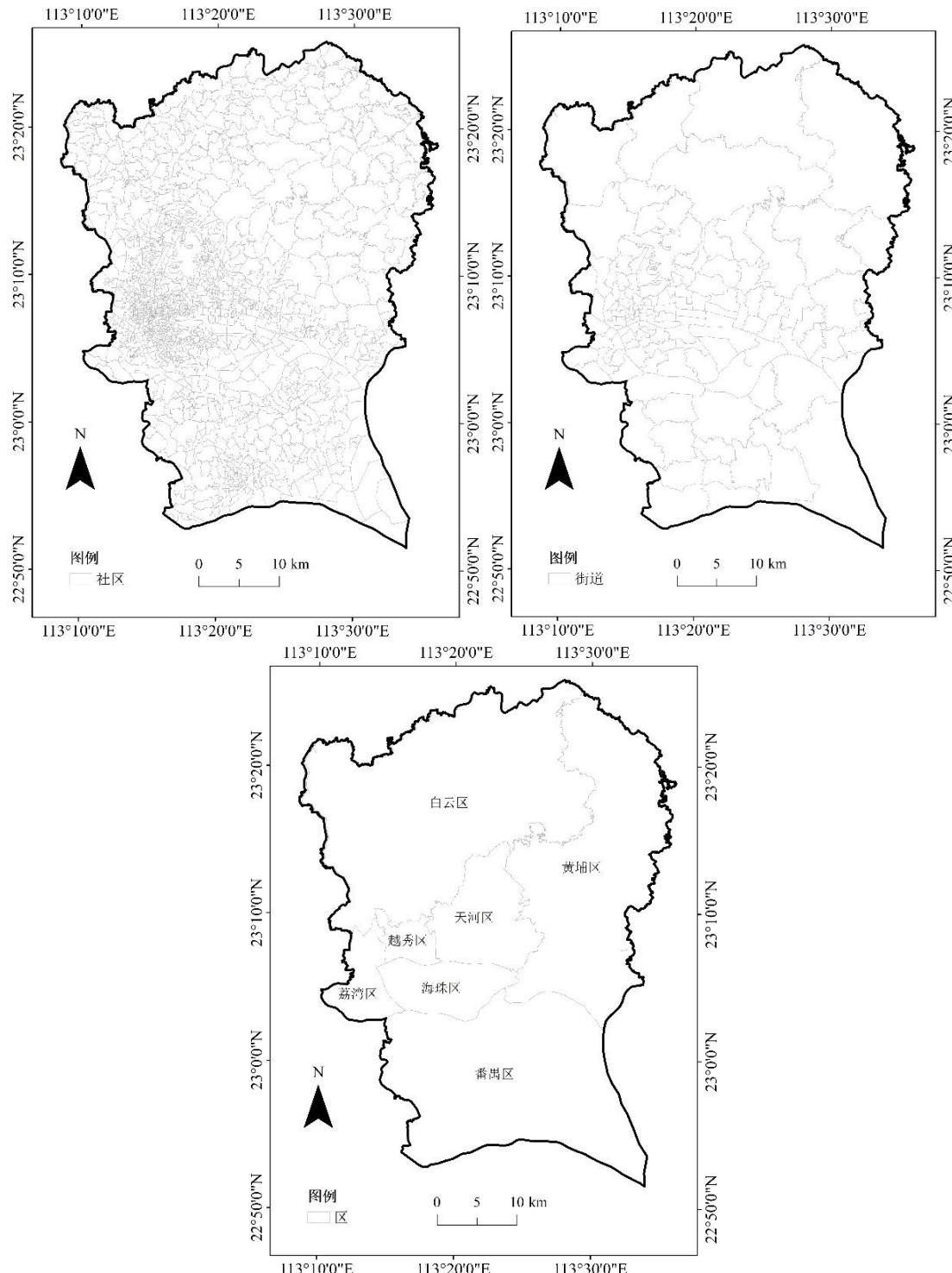


图 2-10 研究区六普人口普查行政单元示意图

本研究使用截止目前可以利用的最新的第六次人口普查数据，数据采集于2010年，简称六普数据。六普数据中包含了不同行政单元下的人口统计数据，包括人口年龄结构、户口来源、性别、受教育程度、房屋情况等各类别信息。基于

社区、街道、县的行政区划如上图 2-10。

#### 2.4.5 POI 设施点数据

相关的 POI 设施点数据为研究团队购买的百度地图道通公司的数据(2016 年更新版), 如图 2-11 所示。包括 15 种大类(涵盖了党政机关、科研教育、医疗卫生、文化体育、金融保险、农林牧渔、商业贸易、住宿餐饮、批发零售、交通运输等各方面), 65 种中类, 共计 451635 个 POI 设施点。

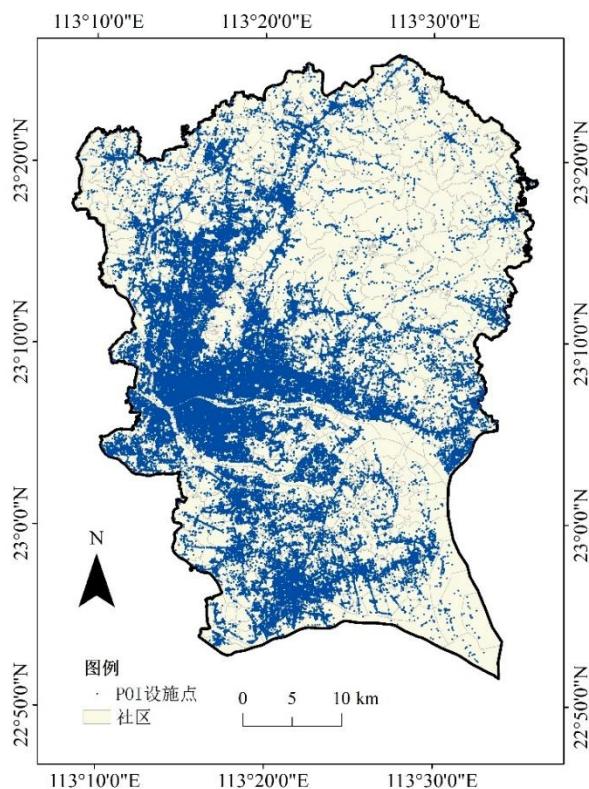


图 2-11 POI 设施点分布

#### 2.4.6 遥感影像数据

遥感影像数据能够充分反应区域的地理环境要素特征。本研究基于美国地质调查局遥感影像数据 Landset8, 获取了研究区 2016 年 1 月的遥感影像数据(图 2-12), 利用 ENVI 完成辐射定标、大气校正等预处理。本研究的影像数据包括红绿蓝三个波段, 该景影像的大小为 11224\*12434 像元。



图 2-12 大气校正后的遥感影像数据

#### 2.4.7 手机基站数据

手机信令数据也是本研究的一种重要数据源。手机信令数据来源于研究团队购买的中国联通公司手机数据。联通手机数据记录日期是 2016 年 12 月 28 日，星期三，为工作日。该手机数据是手机用户的逐小时的轨迹数据，共计有 29.94 亿用户的 306.32 亿个位置记录点。

每个手机用户特定小时的预估位置点是该时段内该用户手机主要时间内最近邻的基站点。举例，晚上 7 点至 8 点的时间段，某用户于甲基站周围活动 30 mins，乙基站周围活动 15 mins，在丙基站周围活动 15 mins。这样，该用户在晚上 7 点至 8 点时间段被记录的活动点就是 A 基站点。

广州市区的基站分布很密集。基站塔的平均间隔距离是 136.4m（Max: 5535.1 m; Min: 0 m; SD: 255.8 m）。本文研究区共有 1856 个人口普查社区，平均面积为  $1.21 \text{ km}^2$ 。因此，基站的主要覆盖场所都被包含在独立的六普社区范围中。从而，通过基站点可以确定用户的估计位置，并通过基站位置和人口普查社区位置的关系可以确定用户每小时所处的人口普查社区单元。

在原始基站数据的基础上，本研究根据基站划分了泰森多边形区域（图 2-13），汇总得到手机数据各项指标。具体估计的方式是，人口密度的估计是通过基站覆盖的面积加权公式得到，社会经济属性的估计是通过对应人口数据的加权公式得到。

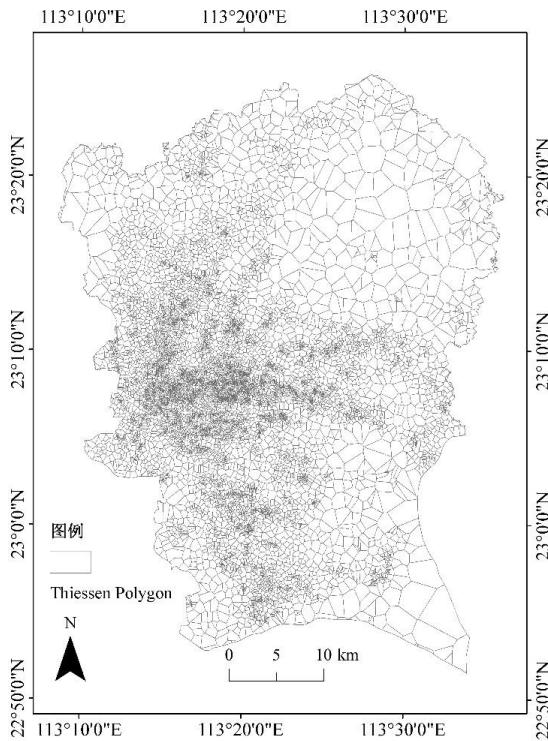


图 2-13 基于泰森多边形的手机基站覆盖范围分布

#### 2.4.8 犯罪案件数据

作为对犯罪情况的感知，犯罪率是与居民安全感知密切相关的因素。研究团队与广州市公安局建立了长期合作协议，每年定期从公安局获取原始接警数据。由于原始接警数据是以文本形式记录，所以需要对于海量接警数据批量地理编码。基于（廖薇薇 et al., 2018）开发的一套地理编码技术，完成犯罪案件数据的空间化。由于本研究的安全感知数据来源于 2016 年初的问卷调查，因此本文获取了研究区 2014 年和 2015 年年度的“两抢两盗”犯罪案件（盗窃犯罪和抢劫犯罪）数据，并根据社区尺度汇总成为总的案件数，基于社区人口密度计算得出犯罪率再除以 2 得到年均社区犯罪率数据。由于犯罪数据在我国属于保密数据，所以不能公开具体的点状数据，图 2-14 展示了基于原始案件点数据的核密度空间化。

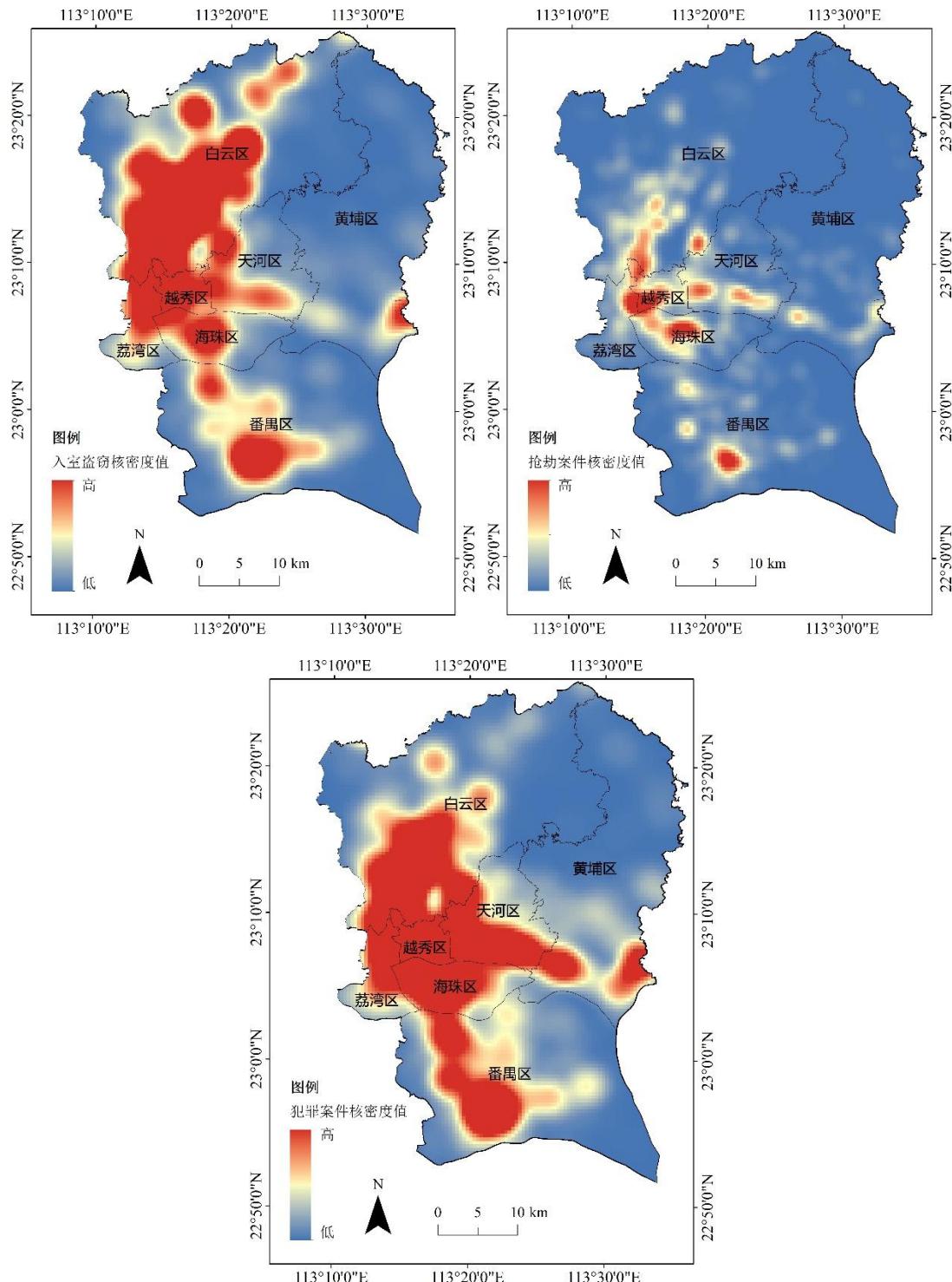


图 2-14 基于核密度的犯罪案件空间分布

## 2.5 小结

本章节详细叙述了研究问题、研究区域、研究框架、研究内容、研究数据等模块。论文确定了四个主要研究问题和五个主要研究内容，一共设计了八个章节的研究框架，包括四个章节的核心研究内容。以国内主要城市之一—广州市七个

中心城区为研究区域。

同时，研究准备了所需要的基础地理信息数据、街景图像数据、手机信令数据、社会经济数据、遥感影响数据、问卷调查数据、犯罪案件数据等。本研究收集的数据集可以为接下来安全感知的环境要素识别、分析，以及安全感知的空间模拟提供充足的数据支撑。同时，本研究数据集也可以为实际的城市规划和管理工作提供精细的数据源。

# 第三章 基于街景图像的安全感知相关微环境要素提取

## 3.1 引言

在城市犯罪和环境设计预防犯罪(CPTED)研究领域,社区建成环境及环境失序信息是预测该地区居民治安安全感的重要因素。诸如社区植被、住宅设计、街道设计、道路交通等建成环境因素都可能影响居民的安全感(张延吉,秦波 & 唐杰, 2017 等)。根据破窗理论(Broken Windows Theory),如果一个社区出现一个破碎的窗户而没有被及时维修,接下来就会有更多的破碎的窗户出现。犯罪分子会认为该社区是实施犯罪活动的良好场所,从而导致犯罪率上升和安全感下降。

关于建成环境的测量,之前研究主要都是通过统计数据或问卷调查的形式获取。这种方法的优点在于准确性强、指标完善,但是存在成本高、主观性强、连续性差等不足。近来,国外一些研究开始采用基于人工打分的街景图像评分等方式来评价建成环境状况(Rundle et al., 2011; Kronkvist, 2014; Quinn et al., 2016)。但是大部分这类研究还只是对小范围、小批量的街景照片进行人工打分,没有进行大规模开展。国内,近年来也有许多研究对街景进行了分析。一方面,从街道要素构成的角度出发,对街道物理环境进行二维或三维评价,包括绿视率、街道建筑物比例、道路比例等。另一方面,从行人使用角度出发,在微观尺度上对细节环境进行评价,比如街道服务效果或街道品质效果的评价。在 Fan Zhang 及其同事开展了一系列研究(Zhang et al., 2018, 2020 & 2021),运用深度学习方法和大规模街景图像数据,分析了北京和上海等大城市中心城区的主客观建成环境要素指标。但是,之前研究主要关注于提取特定建成环境要素,而没有结合安全感知的特定研究需要,专门设计基于环境失序理论的失序微环境体系,也没有实现针对特定城市的大范围环境失序要素制图。

在本章中,基于开源的深度学习模型,本文构建多个对社区环境进行要素识别的模型。本研究主要对街景进行三大方面的处理,包括图像语义分割、图像分类、图像目标检测。这三大方面的任务都有一系列成熟的传统技术方法。在图像语义分割方面,传统上可以通过不同波段或图片处理技术完成提取。图像分类方

面，传统可以通过支持向量机等传统机器学习方法或人工神经网络方法进行处理。在目标检测方面，传统的目标检测方法是从给定图片中选择一个单元，然后从所选单元中提取特征，然后利用常见分类器对获取的特征完成分类。但是，这些传统图像处理方法均存在一些局限性。综合来看，第一，传统方法会通常会有大量重复计算，大批量街景的处理时间较长。第二，传统方法的预测能力不高，而且存在瓶颈，不容易通过传统方法上的改进提升预测能力。深度学习技术的发展给图像处理带来了颠覆性的影响，在图像处理各领域得到广泛利用。卷积神经网络 (Convolutional Neural Networks, CNN)以其局部链接(local connections)、权重分享 (shared weights)的优势使得模型可以高效地学习图片空间的高层次语义信息，进而完成较复杂的任务，在计算机视觉领域受到普遍关注。一系列新的算法被提出，如图像分类的 Inception-v4 算法、ResNet 算法、VGG 算法等，目标检测中的 YOLOv3、Faster-R-CNN、MobileNet-SSD 算法等，图像分割中的 Deeplabv3+、ICNET 算法等。因此，本研究尝试使用一系列先进的深度学习方法，完成图像的分割、分类、目标检测任务，最终得到研究区范围内所有街景采样点的多种街景失序评价数据，作为下一步微环境与安全感知关系研究的重要数据支撑。

本章节的研究流程图如图 3-1。

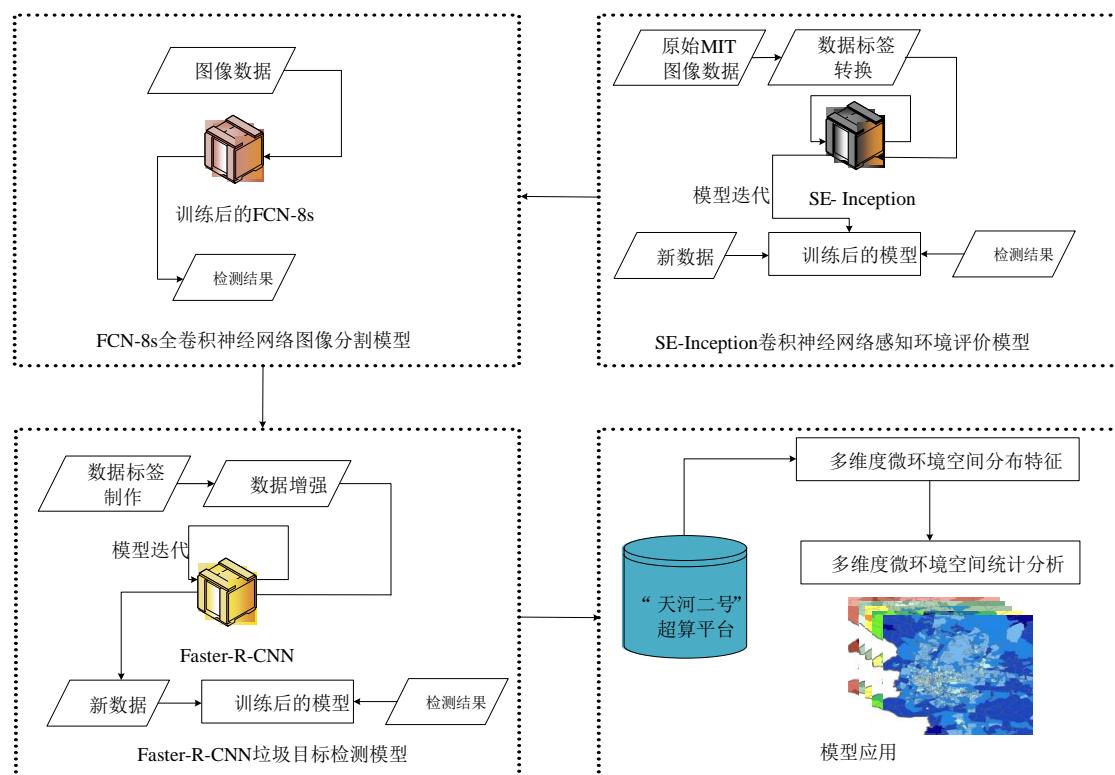


图 3-1 微环境要素提取流程图

## 3.2 基于深度学习技术的微环境要素提取模型实验建模

### 3.2.1 广州超算云计算平台

由于研究区的街景照片总量较大，对大量街景照片进行图像语义分割和目标检测需要高性能的处理平台。本研究使用位于中山大学的国家超算广州中心进行图像后期的实际预测任务。超算平台的主机“天河二号”由国防科技大学研发，位于中山大学广州校区东校园。“天河二号”的计算能力和速度均处于国际一流水平，峰值计算和持续计算速度分别为 10.07 亿亿次/秒和 6.14 亿次/秒，总内存和全局存储分别达到 3PB 和 19PB (1PB = 1024TB)<sup>3</sup>。

超算的主要使用步骤如下。第一步，申请账号，登录中大广州超算中心网站 <https://th2cloud.nscc-gz.cn/project/overview/>。第二步，完成创建云主机操作，主要包括基本信息（名称、数量、类型、规格等）、网络配置、登录设置三个步骤。第三步，配置好云主机，开始工作任务，待结束时再删除云主机。其中，超算处理平台的花费取决于云主机机时、数量、核数等多个参数。

### 3.2.2 基于 SE-Inception 卷积神经网络的感知环境要素模型构建

本研究尝试使用卷积神经网络模型训练街景图片，获得图片美丽度和富裕度两个指标。样本的标注数据来源于麻省理工学院（MIT）媒体实验室 MIT Media Lab 开放的街景打分数据集。对数据集进行处理后，训练基于卷积神经网络的感知环境要素模型，然后利用该模型对研究区内的街景图片进行预测评分，以得到相应的感知维度指标。

#### 3.2.2.1 MIT place pulse 数据集及感知环境要素指标

MIT Media Lab 于 2013 年开展了“Place Pulse 2.0”项目，旨在通过在线数据收集平台收集公众对城市环境的感知评分。在网站上(<http://pulse.media.mit.edu/>)，参与者通过在线游戏的方式，需要在随机给出的两张街景图片中选择一张，选择依据是“哪一个图片更加 X”。其中，X 代表六个维度：安全、无聊、美丽、富裕、

---

<sup>3</sup> <http://www.nscc-gz.cn/Introduction/Generalprofile/Basicsituation.html>

压抑、热闹。参与者需要从三个按钮中选择其中一个，按钮分别为：左边图像、右边图像、相同。图 3-2 显示了参与者打分评价的界面。

该平台的街景照片来自于 28 个国家的 56 个城市，包括了诸如纽约、伦敦、格拉斯哥、哈博罗内等大中小型城市，但城市主要仍集中在欧洲、北美等西方城市。该数据集包括了从 2007 年到 2012 年之间 110988 张街景照片的感知评价得分。评分者来自于 162 个国家的志愿者，多样性强。评分者的人口统计分布特征也很分散，不存在显著相关 (Dubey et al., 2016; Salesses et al., 2013)。此外，用户间的评价结果较为相近，数据集的内部一致性高 (Salesses et al., 2013)。因此，该数据集的质量较高。

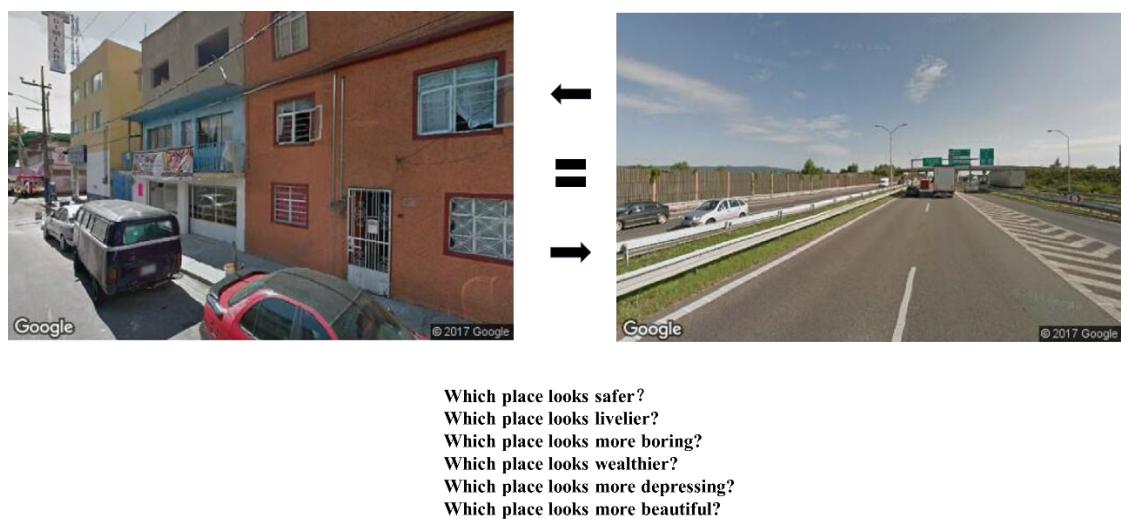


图 3-2 MIT place pulse 数据收集平台（改绘自 Zhang et al., 2018）

### 3.2.2.2 MIT place pluse 数据集感知分数计算

从 Place Pulse 数据可知，评价数据是通过两张图片对比打分获取。虽然直接通过比较的评价数据可以训练模型，但是实际应用中的街景评分大多是给出一个具体数值。因此，本研究将 Place Pulse 数据集的街景评价数据转换成单张街景感知环境评价。转换思路是，一张图片  $i$  可能和其他图片在特定项目比较过若干次，那么  $i$  被选择的次数就表示了人们对特定项目感知的强度。同时，在计算  $i$  的强度的时候，也需要考虑与它做对比的其他图片的强度。

参考之前文献 (Salesses et al., 2013; Ordonez & Berg, 2014)，本文采用进度优势方法(Strength of Schedule Methods)进行评分。对于图像  $i$  的第  $j$  个评价维度，计算公式如下：

第一步，计算积极评价指标( $P_{ij}$ )和消极评价指标( $N_{ij}$ ):

$$P_{i,j} = \frac{p_{i,j}}{p_{i,j} + e_{i,j} + n_{i,j}}, \quad (3-1)$$

$$N_{i,j} = \frac{n_{i,j}}{p_{i,j} + e_{i,j} + n_{i,j}}, \quad (3-2)$$

式中，照片*i*在第 *j* 个评价项与另外照片相比较时，会得到高 ( $p_{i,j}$ )，没有差异 ( $e_{i,j}$ ) 和低 ( $n_{i,j}$ ) 的评估频次。据此，可得出照片 *i* 在第 *j* 个评价项的得分  $Q_{i,j}$ :

$$Q_{i,j} = \frac{1}{3}(P_{i,j} + \frac{1}{p_{i,j}} \sum_{k_1=1}^{p_{i,j}} P_{a_{i,j}^{(k_1)},j} + \frac{1}{n_{i,j}} \sum_{k_2=1}^{n_{i,j}} N_{b_{i,j}^{(k_2)},j}), \quad (3-3)$$

式中，与照片 *i* 对比时被判定与图像 *i* 相比该项得分较低的第 *k<sub>1</sub>* 张照片记为  $a_{i,j}^{(k_1)}$ 。与照片 *i* 对比时被判定与图像 *i* 相比该项得分较高的第 *k<sub>2</sub>* 张图像记为  $b_{i,j}^{(k_2)}$ 。

在 Place Pulse 数据集中，不一定全部图像每个评价项目和另外图像的评价项对比过，所以， $p_{i,j}$  或  $n_{i,j}$  一定概率为 0。因此本研究使用阈值  $M$  删除掉代表性不足的评价值，用 -1 指示空缺值的得分，把公式 (3-3) 修改为：

$$Q_{i,j} = \begin{cases} \frac{1}{3}(P_{i,j} + \frac{1}{p_{i,j}} \sum_{k_1=1}^{p_{i,j}} P_{a_{i,j}^{(k_1)},j} + \frac{1}{n_{i,j}} \sum_{k_2=1}^{n_{i,j}} N_{b_{i,j}^{(k_2)},j}) & \text{if } p_{i,j} + e_{i,j} + n_{i,j} \geq M \\ -1 & \text{otherwise} \end{cases}. \quad (3-4)$$

下图展示了四张街景图像样本的六个维度的评分值。可以看出，第一张图片的安全感和美丽度评价得分比较低，而无聊度和热闹度的评价得分比较高。

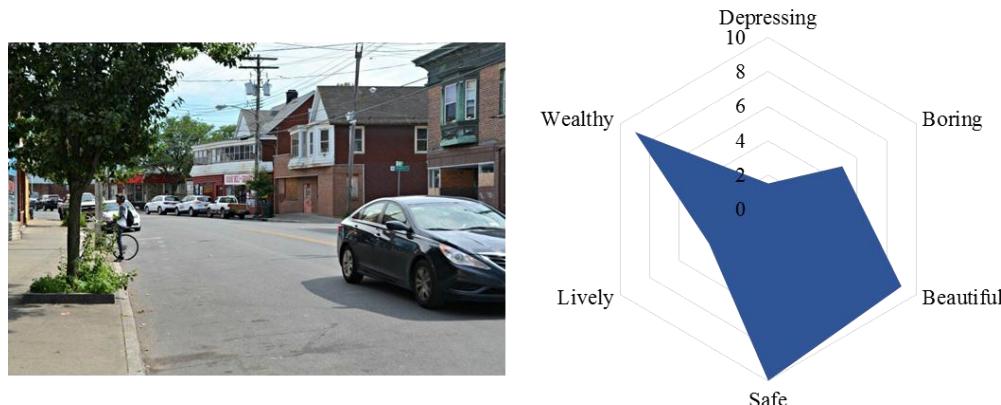


图 3-3 MIT Place Pulse 数据集样本及其感知维度评分示例（样本图片和标签）

### 3.2.2.3 基于 SE-Inception 的感知环境失序要素模型

根据对街景图片的主观评价数据，实现输入单张街景图片可以准确预测出主观评价得分，是一项具有挑战性的任务。本研究基 SE-Inception 模型，完成两种

环境感知失序要素—富裕度和美丽度的度量任务。该模型基于 Multitask Learning, 结合 Inception v3 和 SENet(Squeeze-and-Excitation Networks), 搭建了卷积神经网络模型。

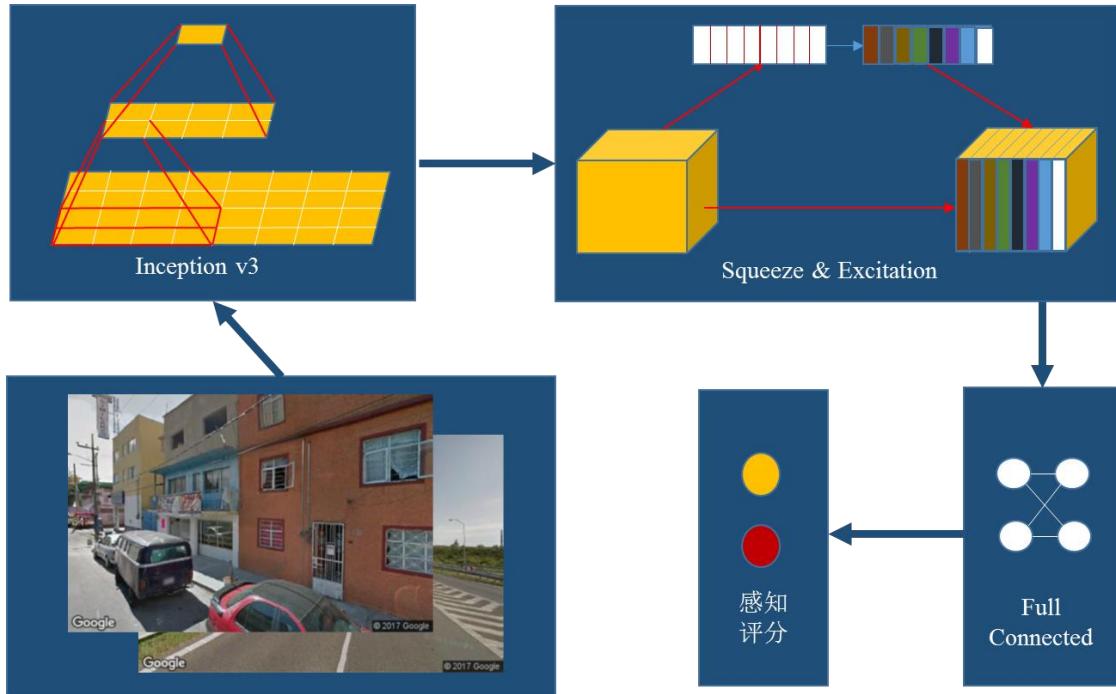


图 3-4 基于 SE-Inception 的 CNN 模型结构图(改绘自 Hu, Shen & Sun, 2018)

其中, RGB 三原色的街景照片最先输入, 它通过 Network Forward Propagation, 得到一个二维向量作为感知失序要素值 (图 3-4)。输入 Place Pluse 数据集的照片集作为训练图像, 将使用进度优势方法(Strength of Schedule Methods)提取的评估值作为训练图像的标签值, 进行模型训练。本研究的损失函数 (Loss Function) 确定为掩膜均方误差 (Masked Mean Squared Error, MMSE)。

$$MMSE = \frac{1}{np} \sum_{i=1}^n \sum_{j=1}^p I_{i,j} (\hat{y}_{i,j} - y_{i,j})^2,$$

$$I_{i,j} = \begin{cases} 0 & \text{if } y_{i,j} = -1 \\ 1 & \text{otherwise} \end{cases} \quad (3-5)$$

公式 3-5 表示,  $n$  是照片数量,  $\hat{y}_{i,j}$  是图片  $i$  第  $j$  个结果变量的估计数据,  $y_{i,j}$  图片  $i$  第  $j$  个结果变量的实际数据。可以看出, Loss 函数的得分越小, 该算法的预测效果越好。

#### 3.2.2.4 模型训练与性能评估

根据 MIT 街景评分照片以及评分计算方法, 本研究对 Place pulse 完成照片

评分工作，得到训练样本图像数据集，使用 SE-Inception 模型进行模型训练。

对于 MIT place pulse 数据集标签制作，罗皓文（2019）发现当阈值  $M=8$  时训练模型可以取得最小的 Loss 函数值。因此，本文选取阈值  $M=8$ （样本个数为 5800）进行模型训练，以保障模型的收敛性与训练样本的有效样本量之间保持平衡。将  $500 \times 500 \times 3$  大小的照片输入 SE-Inception 网络，得到一个 2 维向量。本研究使用 Python 3.6 开发平台，Keras 作为框架，CPU 和 GPU 分别是 Intel Core i7-7700K 和 NVIDIA® V100。

基于评价指标 Loss 函数值对模型效果进行评价。随着模型训练步数的增加，训练集的 Loss 值不断下降，最终 Loss 值为 0.006。验证集的 Loss 值也呈现不断下降的趋势，最终的 Loss 值为 0.007。表明模型的训练和验证结果良好，可以用作下一步的实际预测。

### 3.2.2.5 SE-Inception 模型预测

在得到训练模型的基础上，本文拟对广州市所有街景照片进行评分（流程见图 3-5）。由于数量较大，本研究使用 1 台内存 64g 个人电脑，并行 5 个程序，实际花费约 1 周时间，完成了广州市所有街景照片四个角度的所有照片的美丽度和富裕度评分的提取。

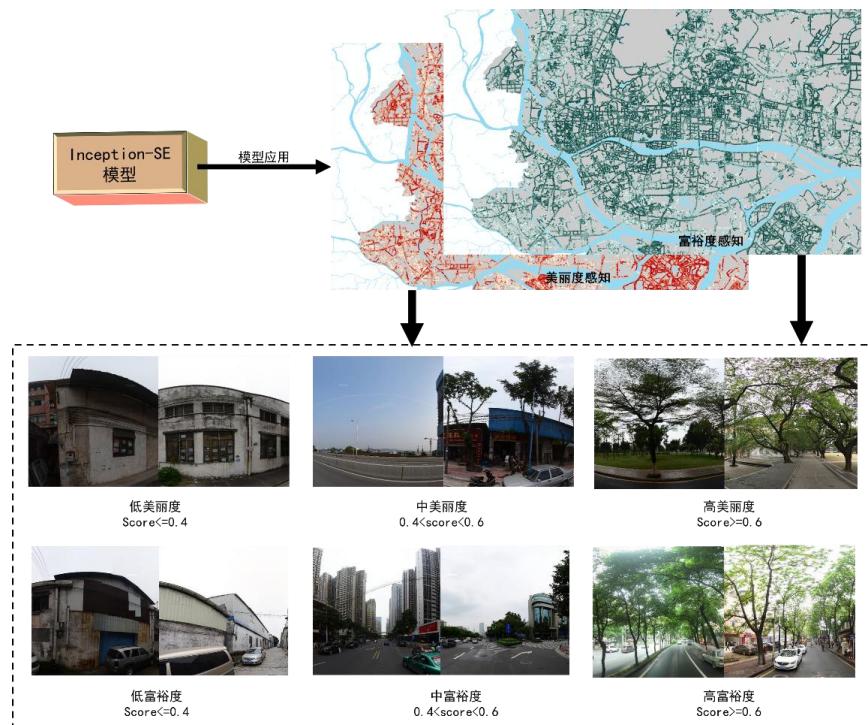


图 3-5 模型预测结果示意图

### 3.2.3 基于 FCN-8s 全卷积神经网络的客观环境要素分割模型构建

为获得街景图像的客观要素指标，本文利用基于全卷积神经网络的图像分割模型，对研究区内的街景图像进行图像分割，得到不同图像要素的占比分值，并从中筛选出与居民安全感知密切相关的绿视率、水视率、天空开敞度、街道路灯比例等指标，进而汇总到空间单元作为环境要素指标。

#### 3.2.3.1 FCN-8s 模型

近年来深度学习方法的发展中，全卷积神经网络可以预测每张图片中每个像素的语义属性，可用于生产自然对象级的分割结果。FCN 模型由 UC Berkeley 的 Long et al. (2015) 开发。FCN 模型与 CNN 模型的最大区别为 FCN 将 CNN 末尾的全连接层转化成卷积层，故得名为全卷积神经网络（Fully convolutional networks，FCN）。本章节使用全卷积神经网络对街景图像完成图像语义分割。

如图 3-6 所示，一个全卷积神经网络 FCN 划分一个街景图片为多个子场景，其中每一个子场景涉及车辆、道路、数目以及其他自然物体，共 150 个种类（不包括一个未知的种类）。在本研究中，使用包括了 150 类样本描述的 MIT ADE20K 数据集<sup>4</sup> (Zhou et al. 2019) 来训练 FCN 网络。以本研究为例，输入是 500\*500 的图像，前五层是卷积层，第五层的输出是 256 个特征图，第 6、7、8 层分别为长度为 4096、4096、150 的一维向量。在 FCN 中第 6、7、8 层都是通过卷积得到的，卷积核的大小全部是 1\*1。

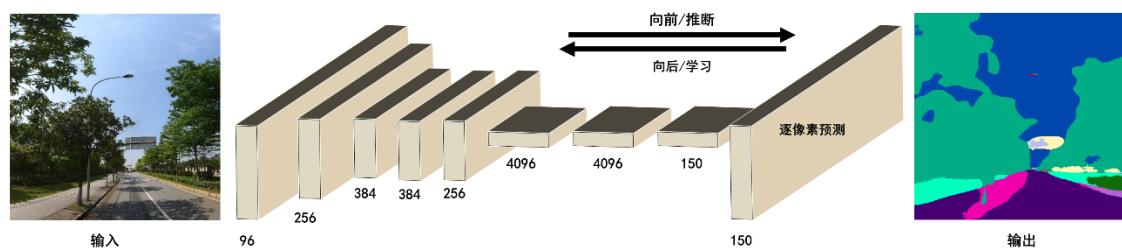


图 3-6 FCN 网络结构图（改绘自 Long et al.,2015）

经过多次卷积后，图片的分辨率不断下降。为了把低分辨率的特征图恢复到原图大小，从而对原图上每一个像素点进行分类预测，需要对特征图进行反卷积，即上采样操作。采取 Long et al.(2015) 的方法，首先对特征图进行一个上池化操

<sup>4</sup> <http://groups.csail.mit.edu/vision/datasets/ADE20K/>

作，再进行反卷积，使得图片分辨率提高到原图大小。在对第 5 层的输出执行 32 倍的反卷积（上采样）基础上，再对第 4 层执行 16 倍的反卷积（上采样）操作，以及对第 3 层进行 8 倍的反卷积操作。因此，这种操作后的网络称之为 FCN-8s 网络。最后，再将这 3 个反卷积的结果图像进行融合。结合基于 softmax 层的逐像素损失计算，FCN-8s 网络通过计算每个分隔掩膜中的像素数量生成图像中每个视觉元素的面积占比（如图 3-7 所示）。

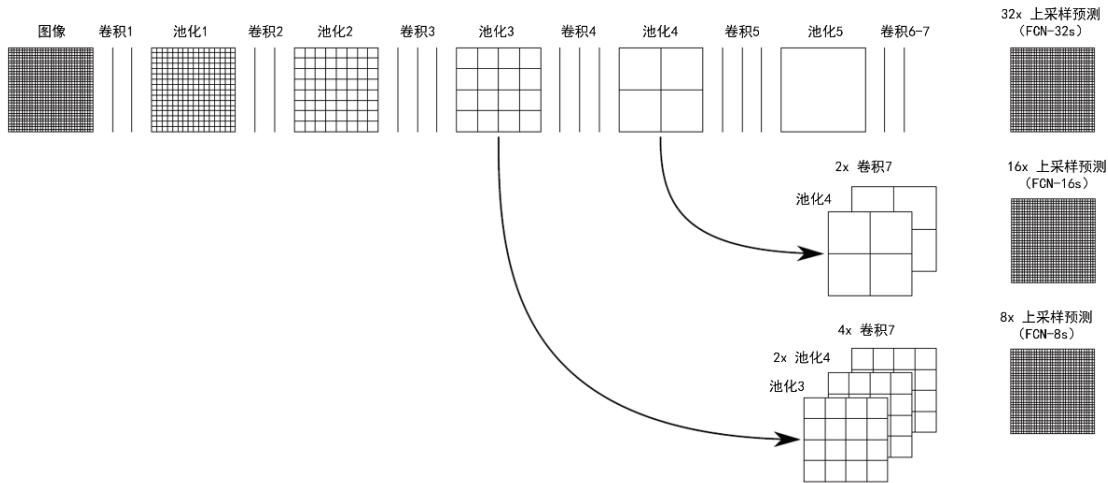


图 3-7 上采样过程图（改绘自 Long et al.,2015）

### 3.2.3.2 街景客观环境要素测度指标

在基于 FCN-8s 完成图像分割后，本文从图像中提取一系列单一要素指标和复合要素指标。具体包括绿视率、水视率、天空开阔程度、建筑物比例以及多样性指标。

#### (1) 绿视率

绿视率是指绿色植被在人眼整个视野范围内的比例。具体公式如下：

$$GVI_n = \frac{G_n}{A_n} = \frac{\sum_{i=1}^4 g_i}{\sum_{i=1}^4 a_i} \quad (3-6)$$

其中  $G_n$  为编号为 n 的照片中各类植被的像素量比例，表示照片 n 的 i 个方向各类植被的像素总值 (East、West、South、North)； $A_n$  为街景照片 n 所有方向的像素的总值。

#### (2) 水视率

$$WVI_n = \frac{\sum_{i=1}^4 w_i}{\sum_{i=1}^4 a_i} \quad (3-7)$$

如式 3-7， $WVI_n$  为编号为 n 的街景照片的水视率（水域可视程度）； $w_i$  为照

片 n 中第 i 个方向的各类水域像素总值;  $a_i$  为街景照片 n 中第 i 个方向的所有像素总值。

(3) 天空开阔程度

$$SVI_n = \frac{\sum_{i=1}^4 v_i}{\sum_{i=1}^4 a_i} \quad (3-8)$$

$SVI_n$  为编号为 n 的街景照片的天空开阔程度;  $v_i$  为照片 n 中第 i 个方向的天空像素总值;  $a_i$  为街景照片 n 中第 i 个方向的所有像素总值。

(4) 建筑物比例

$$BVI_n = \frac{\sum_{i=1}^4 b_i}{\sum_{i=1}^4 a_i} \quad (3-9)$$

$BVI_n$  为编号为 n 的街景照片的建筑物可视程度;  $b_i$  为照片 n 中第 i 个方向的建筑物像素总量;  $a_i$  为街景照片 n 中第 i 个方向的所有像素总值。建筑物比例和人口密度呈正相关。

(5) 路灯可视程度

$$LVI_n = \begin{cases} 0 & \text{if } \frac{\sum_{i=1}^4 l_i}{\sum_{i=1}^4 a_i} = 0 \\ 1 & \text{otherwise} \end{cases} \quad (3-10)$$

$LVI_n$  为编号为 n 的街景照片的路灯可视程度;  $l_i$  为照片 n 中第 i 个方向的路灯像素总量;  $a_i$  为街景照片 n 中第 i 个方向的所有像素总值。当  $LVI_n$  为 1 时, 表明该张照片中有路灯。

(6) 街景多样性

$$DVI_n = \sum_{i=1}^{150} q_i \quad (3-11)$$

$DVI_n$  为编号为 n 的街景照片的多样性水平。 $q_i$  为街景照片 n 中是否有 i 类要素。如果有该类要素, 记为 1, 反之, 记为 0。DVI 的范围为 [0, 150], DVI 越高, 表示该照片中的环境指标类型越多, 更具有丰富性。

### 3.2.3.3 模型训练与评估

本研究使用 Yao et al.(2019)预训练好的网络模型直接进行预测。该网络的训练是通过一组来自 ADE-20K 场景解析和分割数据库的图像(Zhou et al., 2017)完成。ADE-20K 数据集由大量带注释的对象类别(例如树木、植物、车、行人、天空等)组成。该数据集有 20210 项训练数据集和 2000 项验证数据集。在训练 FCN-8s 的过程中, 扫描窗口设置为 500\*500 的像素, 学习率和提前停止的最大小学习率

分别被设置为 0.1 和 0.001。设置 32 为每次输入的 batch size 值。通过两组 Nvidia GTX 1080ti 图形处理器完成模型训练，共计花费时间约 1 周。

训练的 FCN-8s 在自然图像分割上表现良好。使用 ADE-20K 数据集的实验显示，通过逐像素的比较，训练的 FCN-8s 在训练数据集实现了 81.44% 的准确性，以及在测试数据集的准确性为 66.83%。

### 3.2.3.4 FCN-8s 模型预测

利用中山大学“天河二号”超算平台，开通 4 台虚拟机，配置为 12 核 32GB 内存并带有 1 张 v100GPU 显卡（相比于 GTX1050 和 GTX 1080ti 等显卡，处理速度更快）的 windows 10 操作系统，实际花费约 2 周时间，完成了图像分割的运行。

将街景图像输入训练过的网络得到图像分割后，确定植被(如树、草、植物、棕榈树)、水体(如河流、湖泊、喷泉、瀑布、游泳池)、建筑物等的比例。

具体的模型预测应用流程见图 3-8。

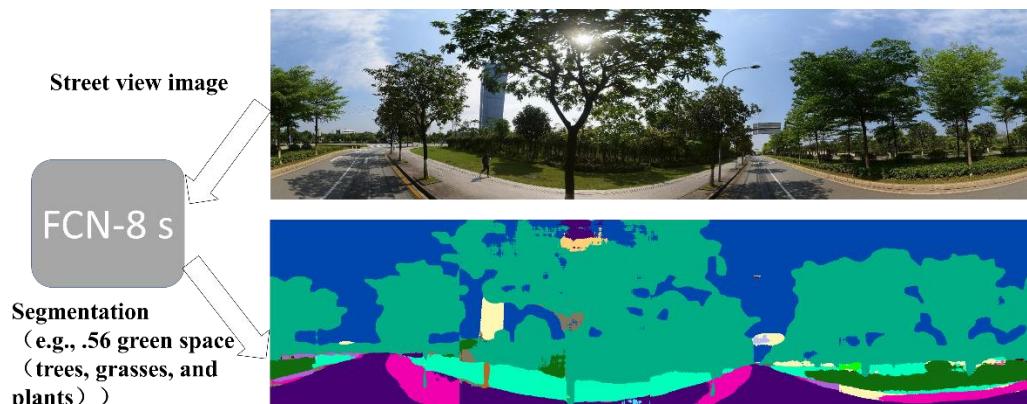


图 3-8 全卷积神经网络工作流程图

### 3.2.4 基于 Faster-R-CNN 卷积神经网络的垃圾目标检测算法

Faster-R-CNN (Faster Region-based Convolutional Neural Network) 方法，即高速区域卷积神经网络方法，是一种被广泛使用的目标检测算法，由 Girshick (2015)最早提出。这个方法的优点是检测速度快、实现自动提取、泛化功能强等。

### 3.2.4.1 训练数据集及数据标注

由于街道垃圾多种多样，很难实现统一的标准进行分类提取。因此，为了简化模型，提升模型的预测准确度，本研究选取建筑垃圾为垃圾识别客体。

由于目前网络上没有公开的街景垃圾数据集，本研究制作专门的数据集和数据标签，作为模型的训练数据。数据集的制作有 3 个步骤。

第一步，从采集到的街景照片中人工筛选出 236 张含建筑垃圾照片，筛选条件是照片中需要有建筑垃圾（如图 3-9）。

第二步，使用在计算机视觉领域广泛运用的 Labelme 标注程序对样本进行标注，框选出包含建筑垃圾的矩形区域，该程序会自动生成 XML 格式的数据，作为对应街景图片的标签数据（如图 3-9）。



图 3-9 带有垃圾的街景照片标注示例

第三步，进行数据增强。由于人工筛选的垃圾照片数量较少，为了提升模型的预测精度，本文使用了一系列图像增强手段对样本图片进行处理，处理方式包括裁剪、平移、改变亮度、高斯模糊、cutout 处理（随机将样本中的部分区域减掉，并且填充 0 像素值）。每一张样本图像，随机选择一种数据增强方式处理。经过十倍的数据扩充，数据增强后的样本达到 2360 张。

### 3.2.4.2 Faster-R-CNN 模型

相较于传统的 Fast-R-CNN 模型，Faster-R-CNN 模型补充了 RPN 框成立算

法，目标的检测可以更加快捷。

Faster-R-CNN 模型由以下几部分组成：

- (1) 图像输入的训练数据集；
- (2) 卷积层( Conv layers )作为基础网络，以提取得到特征图( feature maps )。Faster-R-CNN 使用一组 conv+relu+pooling 层提取图像的特征图。特征图用于之后的 RPN 层和连接层。
- (3) RPN( Region Proposal Networks )层。通过 RPN 网络生成 region proposals。这一层以 softmax 判断 anchors 属后项 (foreground) 或后项 (background)，再利用 bounding box regression 方法调试 anchors，获取准确的 proposals。
- (4) RoI Pooling。RoI Pooling 接收输入的特征图以及 proposals，然后捕捉出 proposal feature maps，之后用于全连接层进行垃圾的判别。
- (5) Classification。使用 proposal feature maps 计算 proposals 的类别，然后再次利用 bounding box regression 方法获取检测框的精确位置。

基本框架如下图 3-10 所示。

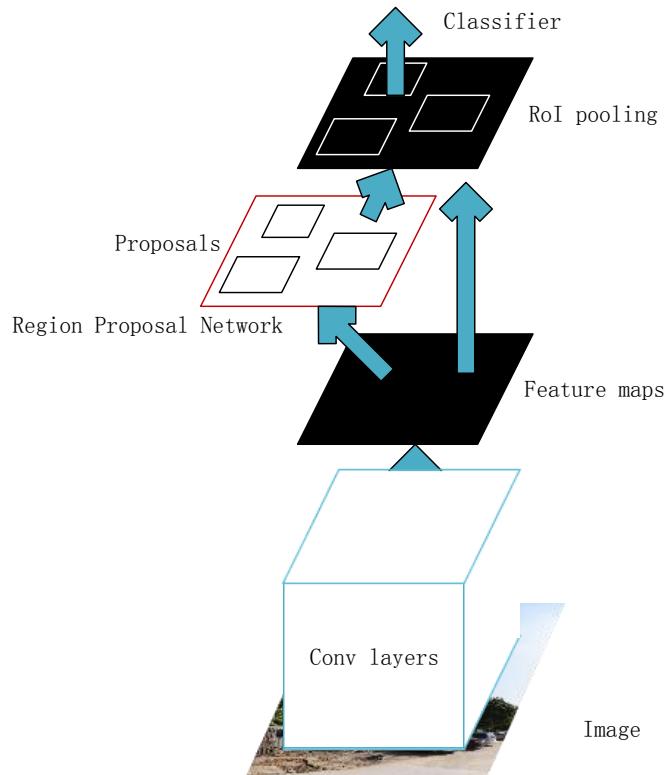


图 3-10 Faster-R-CNN 基本框架 (自绘，改自 Ren et al., 2015 & 2016)

本研究使用 Keras 作为框架，基于 VGG18 模型搭建 Faster-R-CNN 模型。该模型可以输入任意大小的图片。

### 3.2.4.3 垃圾度指标

根据训练模型，如果照片中的一个角度识别出有建筑垃圾，那么该角度的垃圾度记为 1 分，如果照片中一个角度没有识别出任何垃圾，那么该照片的垃圾度记为 0 分。总体垃圾度计算公式为：

$$GAVI_n = \sum_{i=1}^4 q_i \quad (3-12)$$

$GAVI_n$  为街景点 n 的垃圾度值，表示街景点 n 的 i 个方位（东、西、南、北）的照片的垃圾度值。以求和的形式，汇总四个角度的值为该街景点的垃圾度。

### 3.2.4.4 模型训练与评估

本研究使用 Python 3.6 为开发环境，Keras 作为框架，搭建 Faster-RCNN。在准备好训练数据后，图片尺寸均为  $500\times500$  像素，使用 VGG 网络作为网络的基本结构，通过不断迭代训练模型中的多层神经网络，可以自动学习到建筑垃圾的相关特征，并通过反向传播机制相应调整模型中涉及的众多参数。在参数设置上，Faster-R-CNN 包括了 RPN 和分类器两个训练网络，其中优化器参数选择的都是 Adam，学习率为  $1e-5$ 。总网络使用 sgd 优化器用来降低损失函数(Loss)，优化器参数选择为 keras 默认参数，最终的 loss 函数选取 MAE。选择 0.7 为评判为建筑垃圾的概率值。在 GPU 配置的主机上，共花费约 3 天时间完成模型训练。

随机选取照片构成测试集，进行了 5 次实验，以多次实验结果的均值和标准差作为最后的结果。如表 3-1，模型的平均检测精度达到 88.14%，标准偏差为 0.06，在 NVIDIA V100 GPU 上的检测速度达到每张 1 秒。通过测试验证，该模型达到的检测精度基本满足了街道建筑垃圾的检测要求。

表 3-1 垃圾目标检测实验结果

测试图片数量 (张)	检测速度 (秒/张)	检测效果	
		平均检测精度 (%)	检测精度标准差
50	1	88.14	0.06

### 3.2.4.5 Faster-R-CNN 模型预测

具体的模型预测应用如下图 3-11 所示。输入图片后，识别为垃圾的对象在矩形框中被标注，矩形框上方的数值表示算法对该区域确定是垃圾的置信度（以百

分比显示，省略了小数位数）。示例图显示，对于不同街道场景不同形状的建筑垃圾，本研究训练的 Faster-R-CNN 都可以准确的识别，而且识别的准确性较高。

利用中山大学“天河二号”超算平台，开通两台虚拟机，配置为 12 核 32GB 内存并带有 1 张 v100 显卡的 windows 10 操作系统，实际花费约 3 周时间，完成了研究区照片目标检测的运行。



图 3-11 Faster-R-CNN 图像垃圾检测示例

### 3.3 广州市中心城区环境要素的描述统计及可视化分析

#### 3.3.1 感知微环境要素图像评价结果

图 3-12 展示了所有街景点的感知微环境要素值的频率分布，富裕度和美丽度频率高的得分值都集中在 0.45-0.55 的区间，小于 0.4 和大于 0.6 的值分布都较少。表 3-2 展示不同行政区域感知微环境要素的结果。其中，美丽度最高的区是越秀区和黄浦区，最低的区是荔湾区和白云区；富裕度最高的区是海珠区和越秀区，最低的区是番禺区和白云区。

如图 3-13 和图 3-14 分别为两种尺度的感知建成环境指标(富裕度和美丽度)空间分布的结果。可以看出，美丽和富裕感知较高的区域重合较多。美丽和富裕的区域主要集中在以越秀和荔湾为代表的老城区，以及各大高校的集中区域(如中大南校区、广州大学城等区域)、商务区(环市中路、珠江新城等)、高端住宅区(二沙岛、华侨新村、白云堡等)、公园区域。而白云区北部、海珠区南部、番禺区南部等区域的美丽度和富裕度较低。这些区域主要是批发市场集聚区、城乡混合区域。

表 3-2 不同行政区域感知微环境要素结果统计

## 要素

	美丽度		富裕度	
	Mean	Stdev.	Mean	Stdev.
海珠区	0.485	0.022	0.522	0.022
番禺区	0.483	0.02	0.505	0.017
黄浦区	0.495	0.019	0.516	0.019
荔湾区	0.478	0.073	0.517	0.08
天河区	0.49	0.021	0.522	0.02
越秀区	0.494	0.023	0.536	0.022
白云区	0.478	0.022	0.506	0.019

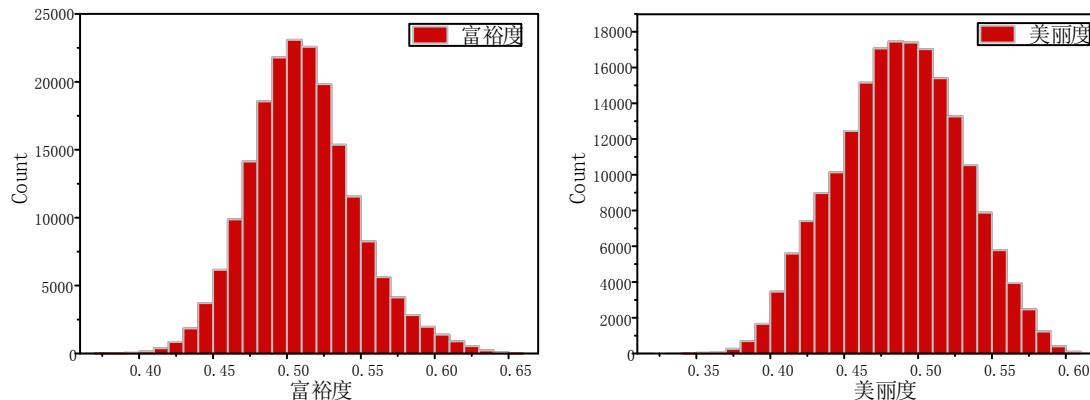


图 3-12 感知微环境要素频率分布直方图

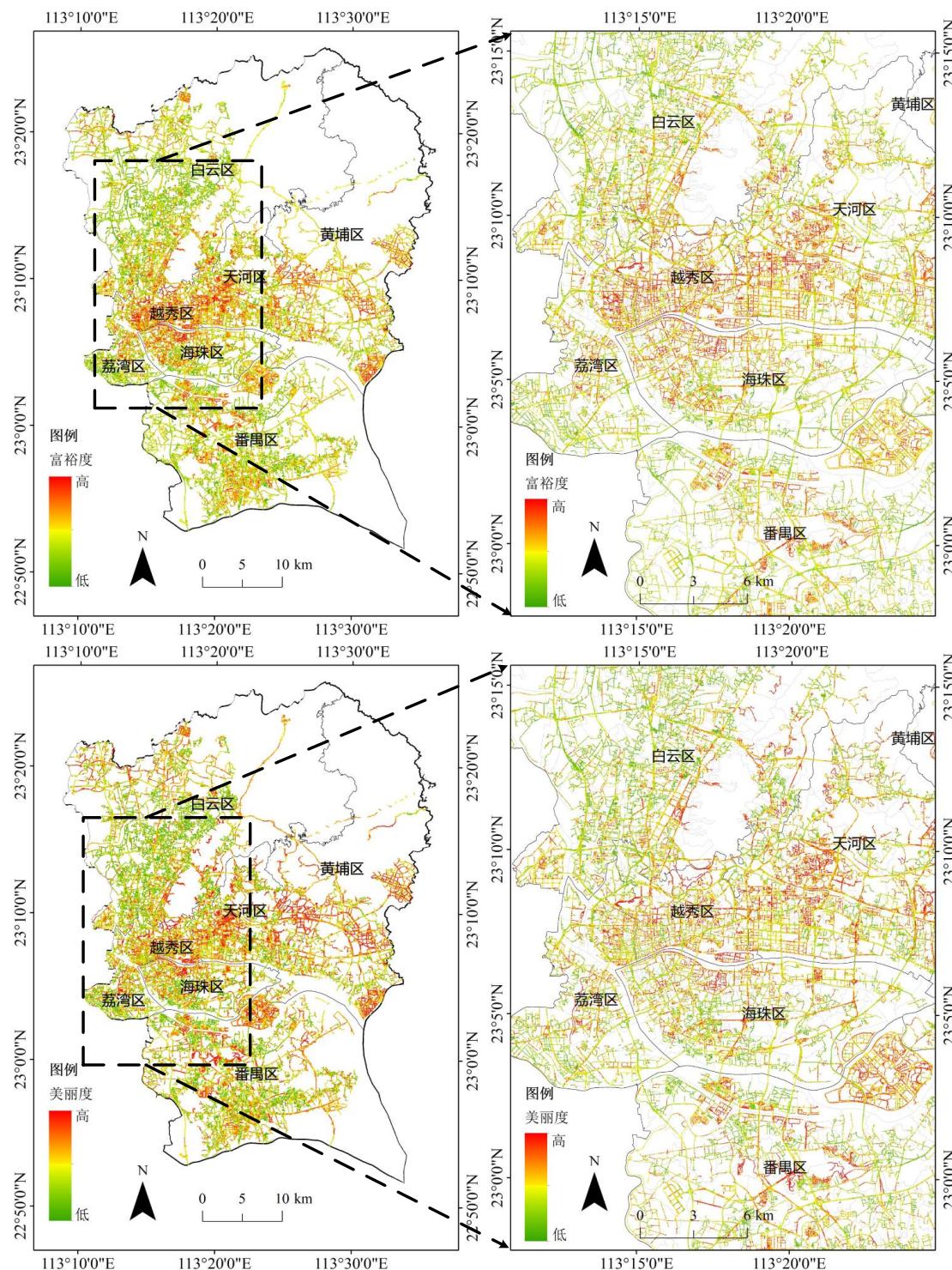


图 3-13 样点尺度下感知微环境要素空间分布结果

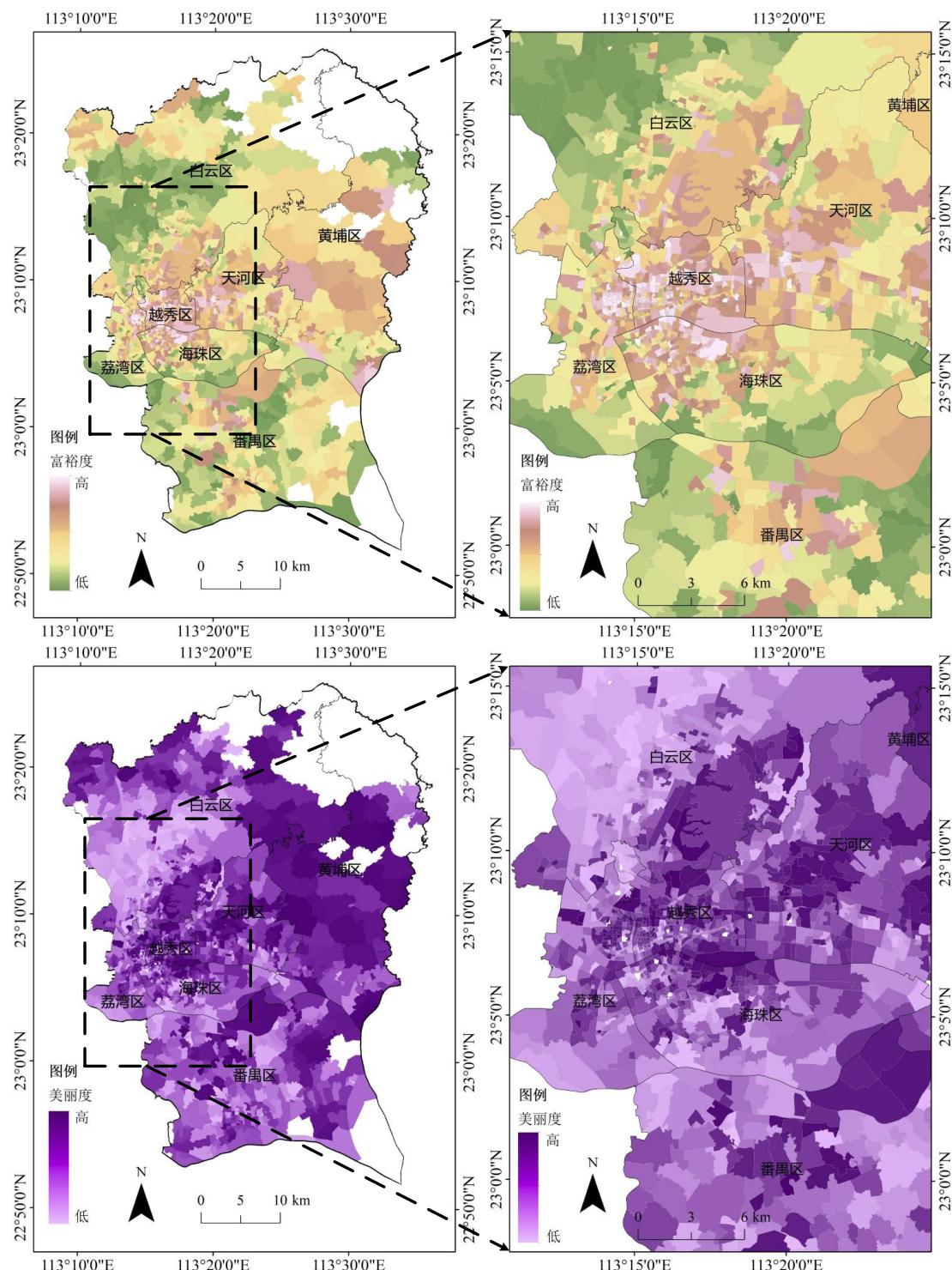


图 3-14 社区尺度下感知微环境要素空间分布结果

### 3.3.2 客观微环境要素图像分割结果

图 3-15 ~ 图 3-17 展示了社区尺度下的客观微环境指标空间分布结果（由于篇幅问题，点状尺度下的客观微环境分布结果见附表 1）。可以看出，绿视率较高的区域集中在主城区、番禺大学城以及天河区北部。水视率较高的区域分布很少，主要分布于珠江水系的流经区域。天空开阔程度指数的空间分异很明显，主城区等城市化水平较高的区域的天空可视率较低，而番禺、白云、天河等区域的外围部分天空可视率较高。建筑物比例指标的空间分布格局与天空开阔指数相反，高值区域集中在主城区商业中心区、人口密集区域。路灯比例的密集区域较为分散，没有相对明显的空间分异特征。多样性指标的高值区域集中在主城区，而外围郊区的多样性值较低。说明广州市的主城区的城市景观较为丰富，而郊区较为单一。

图 3-18 展示了所有街景点的客观微环境要素值的频率分布，表 3-3 展示不同行政区域微环境要素分布的结果。其中，路灯比例最高的区是番禺区，最低的区是越秀区。水视率各区的水平都较低。绿视率最高的区是越秀区，最低的区是白云区。天空比例最高的区是黄浦区，最低的区是越秀区。多样性最高的区是荔湾区，最低的区是黄埔区。建筑物比例最高的区是越秀区，最低的区域是黄浦区。

表 3-3 不同行政区域客观微环境要素结果统计

要素	统计指标	海珠区	番禺区	黄浦区	荔湾区	天河区	越秀区	白云区
路灯比例	Mean	0.562	0.617	0.609	0.551	0.591	0.532	0.542
	Stdev.	0.147	0.126	0.115	0.161	0.109	0.158	0.118
水视率	Mean	0.001	0.002	0.002	0.001	0.001	0.001	0.002
	Stdev.	0.001	0.001	0.004	0.001	0.001	0.001	0.001
绿视率	Mean	0.218	0.207	0.228	0.232	0.243	0.25	0.193
	Stdev.	0.07	0.067	0.066	0.085	0.066	0.077	0.064
天空比例	Mean	0.276	0.373	0.388	0.251	0.297	0.226	0.337
	Stdev.	0.068	0.093	0.076	0.09	0.061	0.058	0.087
多样性	Mean	25.505	24.133	23.549	26.014	24.598	25.964	24.752
	Stdev.	5.701	4.97	4.624	5.307	4.977	5.349	5.131
建筑物 比例	Mean	0.23	0.168	0.123	0.236	0.195	0.249	0.208
	Stdev.	0.072	0.08	0.063	0.082	0.065	0.072	0.083

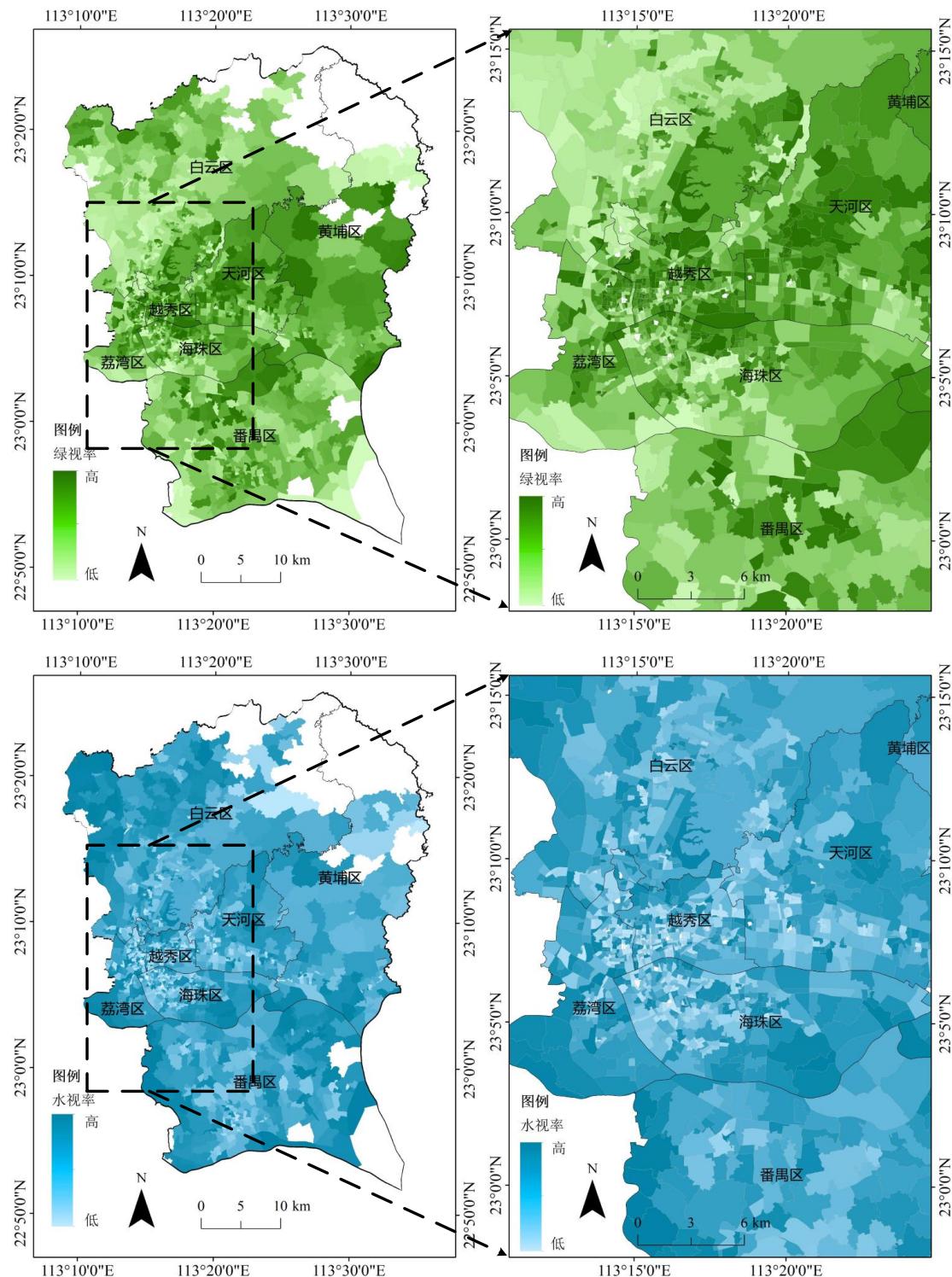


图 3-15 社区尺度下绿视率和水视率微环境要素空间分布结果

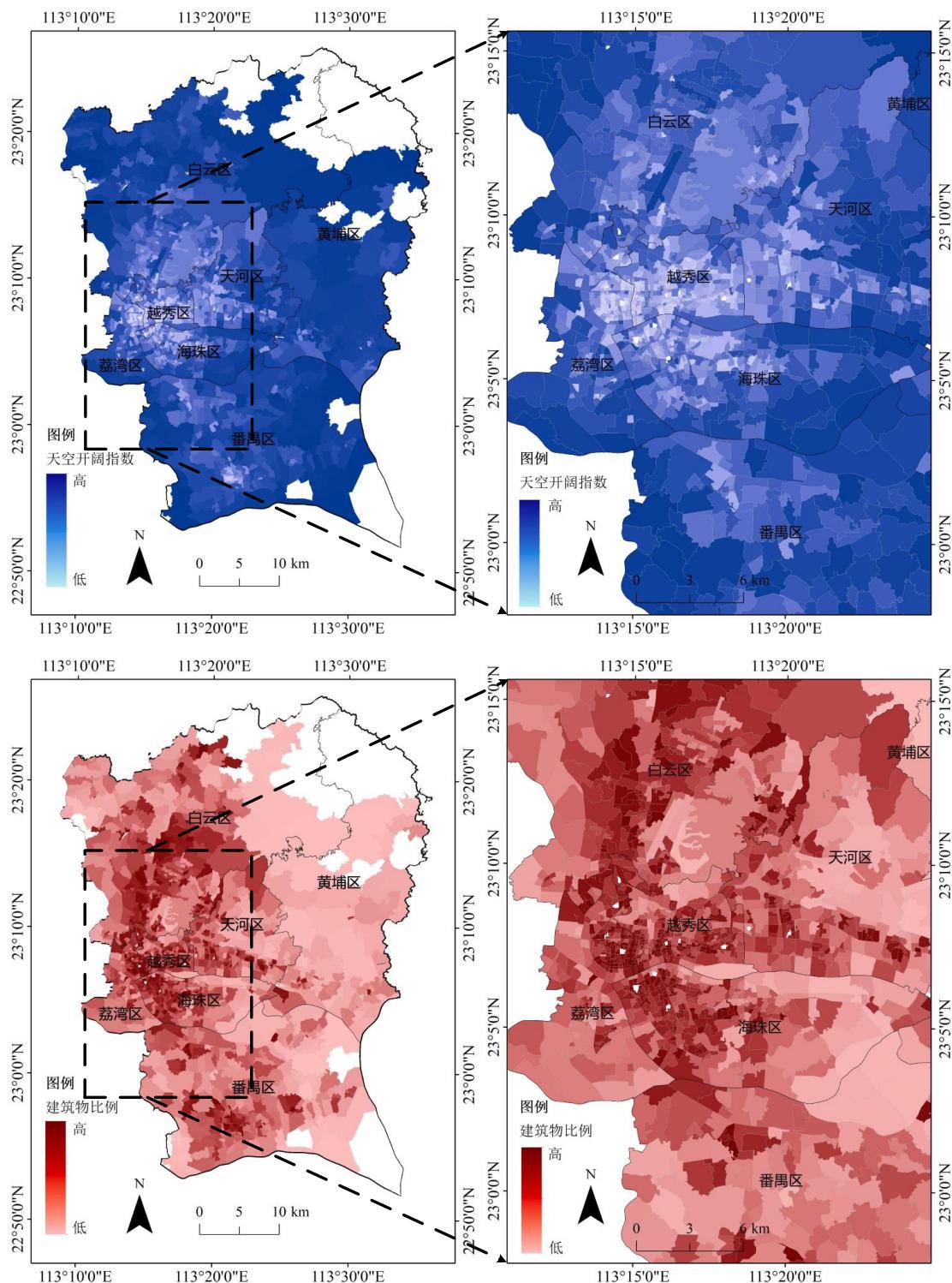


图 3-16 社区尺度下天空开阔程度和建筑物比例微环境要素空间分布结果

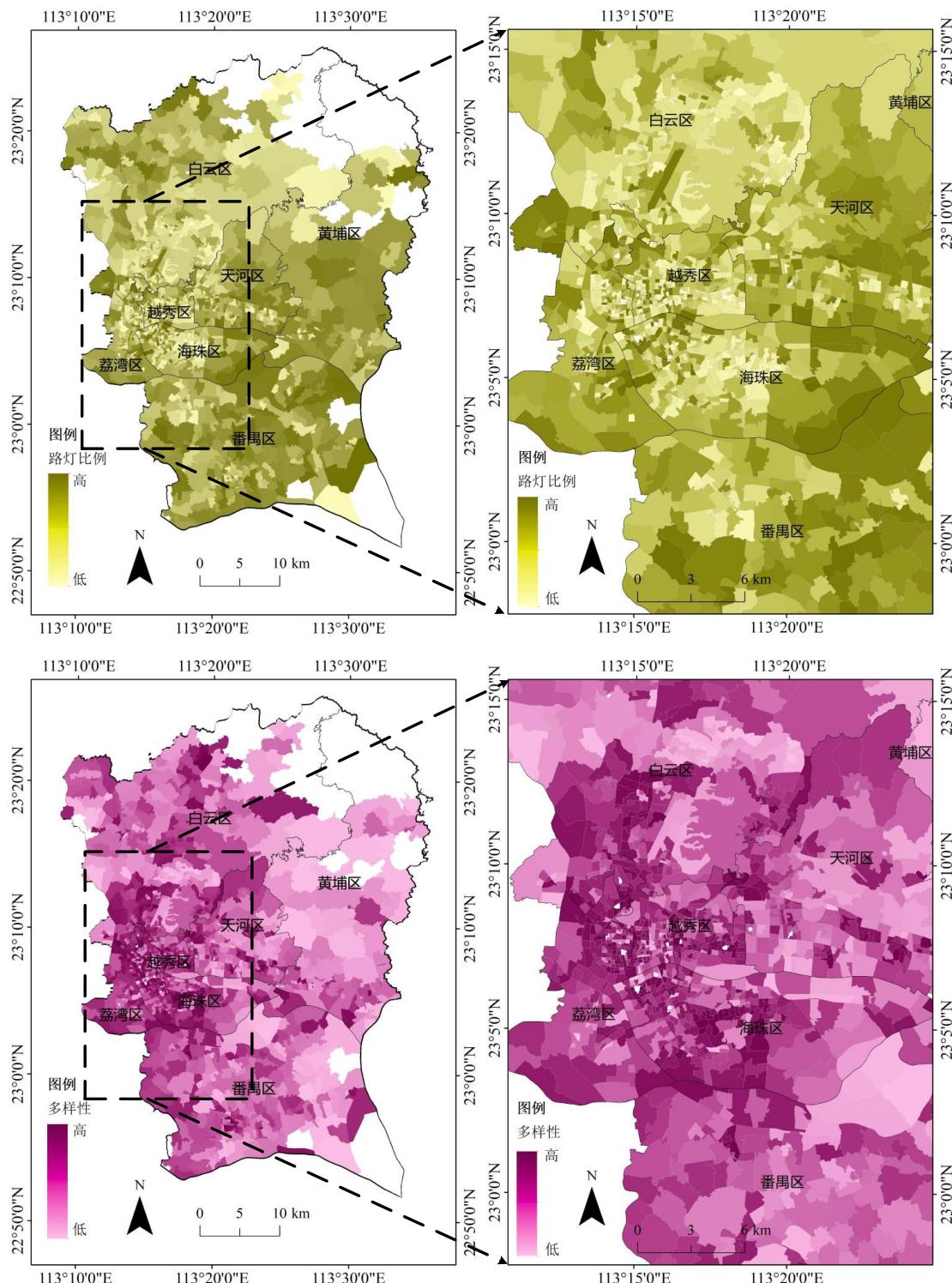


图 3-17 社区尺度下路灯比例和多样性微环境要素空间分布结果

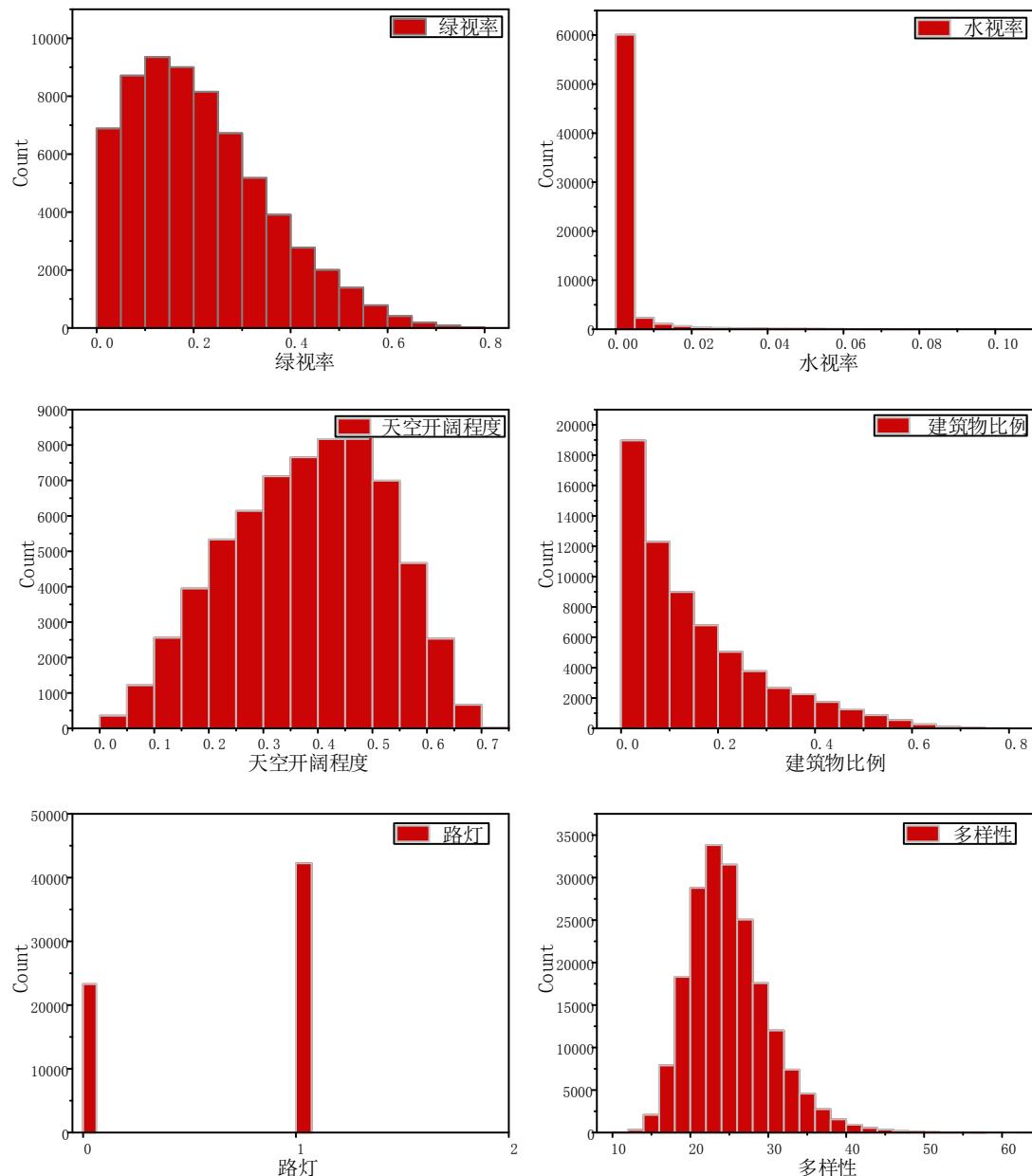


图 3-18 客观微环境要素频率分布直方图

### 3.3.3 垃圾微环境要素目标检测结果

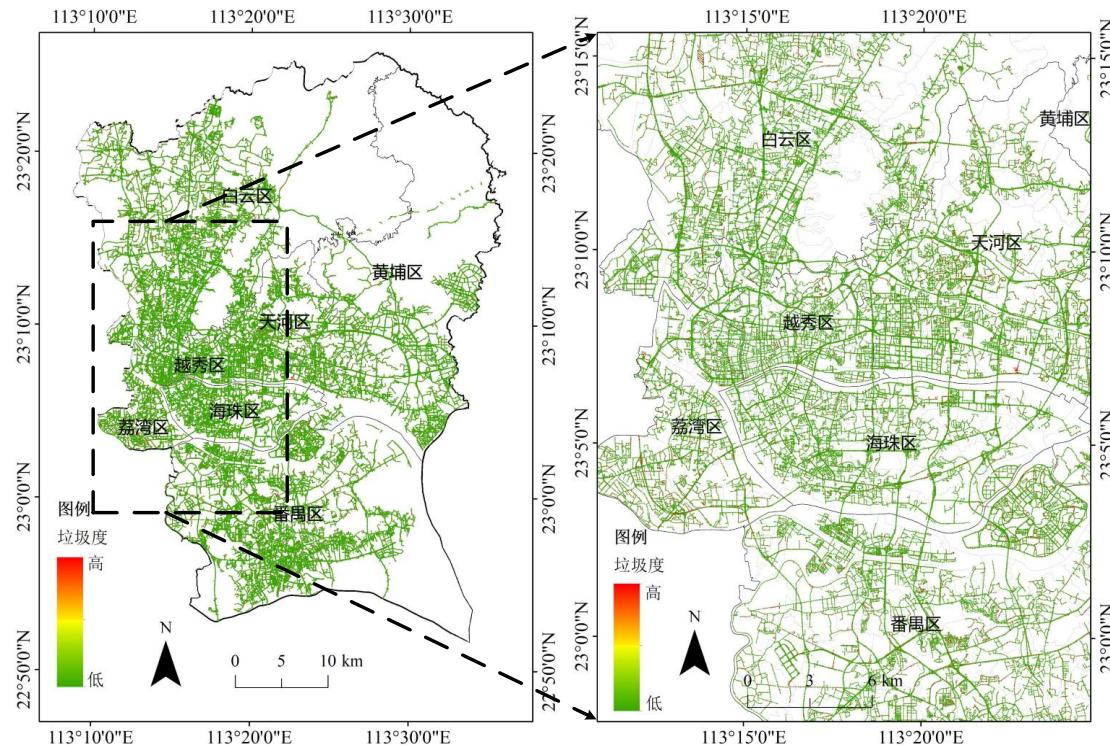


图 3-19 垃圾要素空间分布结果

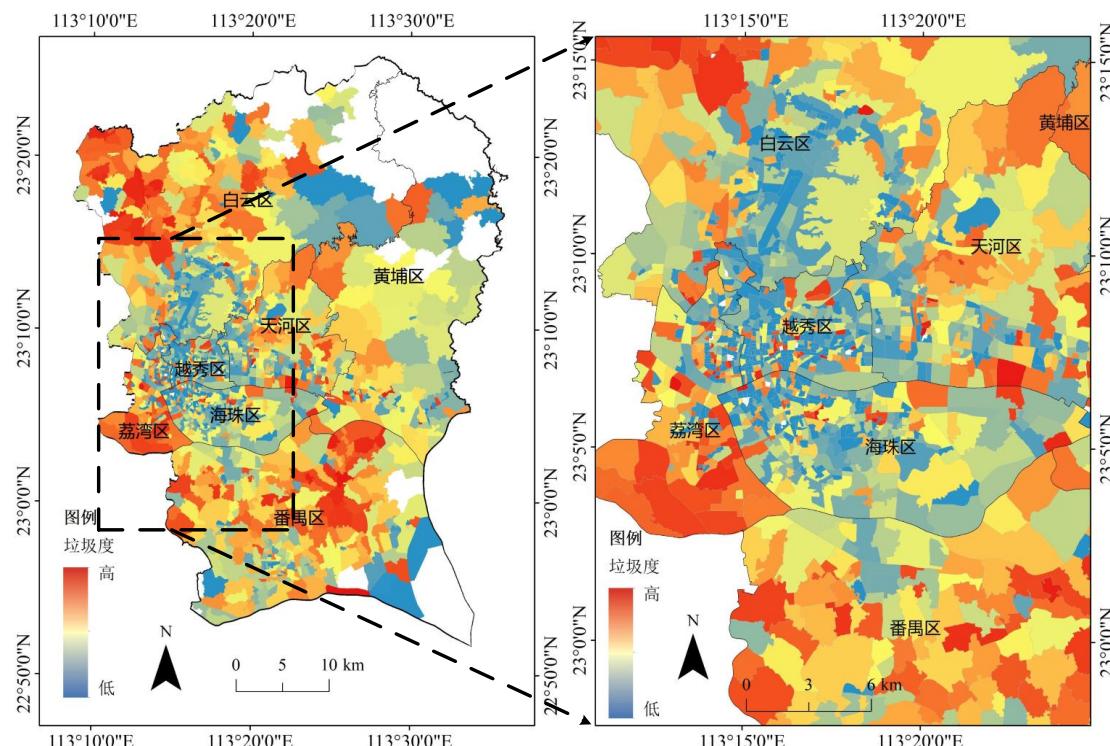


图 3-20 社区尺度的垃圾要素空间分布结果

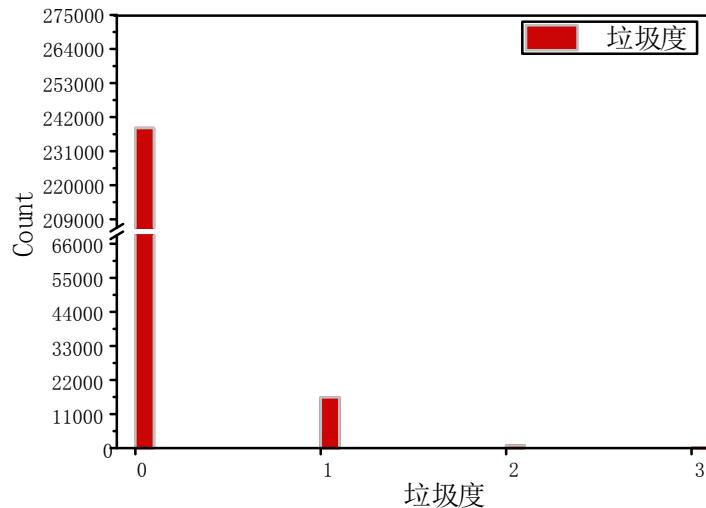


图 3-21 垃圾要素频率分布直方图

表 3-4 不同行政区域客观建筑垃圾要素结果统计

要素	统计指标	海珠区	番禺区	黄浦区	荔湾区	天河区	越秀区	白云区
垃圾度	Mean	0.048	0.075	0.064	0.09	0.065	0.04	0.074
	Stdev.	0.021	0.02	0.015	0.016	0.221	0.216	0.019

建筑垃圾的分布如图 3-19、图 3-20、图 3-21 和表 3-4 所示。可知，建筑垃圾的整体分布较少，大部分街景样本点没有建筑垃圾，说明广州整体城市环境较为干净有序。社区尺度来看，一部分社区没有检测到有建筑垃圾的分布，这些社区大部分分布在主城区，如越秀区、海珠区北部、天河区南部，以及白云区南部区域。由此可知，广州市主城区城市环境较好，建设施工较少，这也与现实认知基本相符。建筑垃圾分布相对的高值区域主要在荔湾区、白云区、番禺区。其中，白云区和番禺区是近年发展较快的区域，因此建筑施工区域可能较多，导致建筑垃圾街景检测值偏高。而荔湾区可能是由于旧城改造和更新，导致建筑垃圾水平相对较高。

### 3.3.4 相关分析与结果验证

从描述统计和可视化分析中可以发现，其中一些街景微环境因素的空间分布具有相似性。因此，本研究进一步使用斯皮尔曼相关系数对点状的微环境要素值进行了相关分析。结果如表 3-5 所示。其中，绿视率和美丽度、富裕度得分的相关性较高，呈正相关。绿视率得分越高，街景的美丽度和富裕度得分越高。建筑物比例和天空比例、美丽度的相关性较高（大于 0.5）。建筑物比例越高，天空比例越小，街景美丽度得分也越小。美丽度和富裕度之间呈高度的正相关，相关

系数  $\alpha=0.88$ 。其他微环境要素之间的相关性程度相对较低。

表 3-5 相关分析结果表

类别	水视率	绿视率	天空比例	建筑物 比例	多样性	美丽度	富裕度	垃圾度
水视率	1	0.013	0.006	-0.059	0.135	0.009	-0.035	0.257
绿视率	0.013	1	-0.417	-0.398	-0.189	0.744	0.659	-0.07
天空比例	0.006	-0.417	1	-0.517	-0.303	-0.107	-0.321	0.263
建筑物 比例	-0.059	-0.398	-0.517	1	0.281	-0.545	-0.272	-0.134
多样性	0.135	-0.189	-0.303	0.281	1	-0.244	-0.123	-0.06
美丽度	0.009	0.744	-0.107	-0.545	-0.244	1	0.88	-0.007
富裕度	-0.035	0.659	-0.321	-0.272	-0.123	0.88	1	-0.123
垃圾度	0.257	-0.07	0.263	-0.134	-0.06	-0.007	-0.123	1

此外，本研究对基于街景图像挖掘的特定指标进行了简单验证。基于遥感影像，本研究分别计算了研究区社区尺度的 NDVI 值和 NDWI 值，对基于街景图像提取的绿视率和水视率进行初步验证（图 3-22、图 3-23）。目视分析来看，基于遥感的归一化植被指数 NDVI 值和绿视率的空间分布较为吻合（相关系数  $\alpha=0.17$ ），老城区（荔湾区、越秀区）以及郊区（天河区北部、番禺区北部）都是植被绿化较为密集的区域。基于遥感的归一化水体指数 NDWI 和水视率的空间分布的相关性较小（相关系数  $\alpha=0.05$ ）。不过，NDWI 和水视率的高值都相对集中在珠江水系的流经区域。

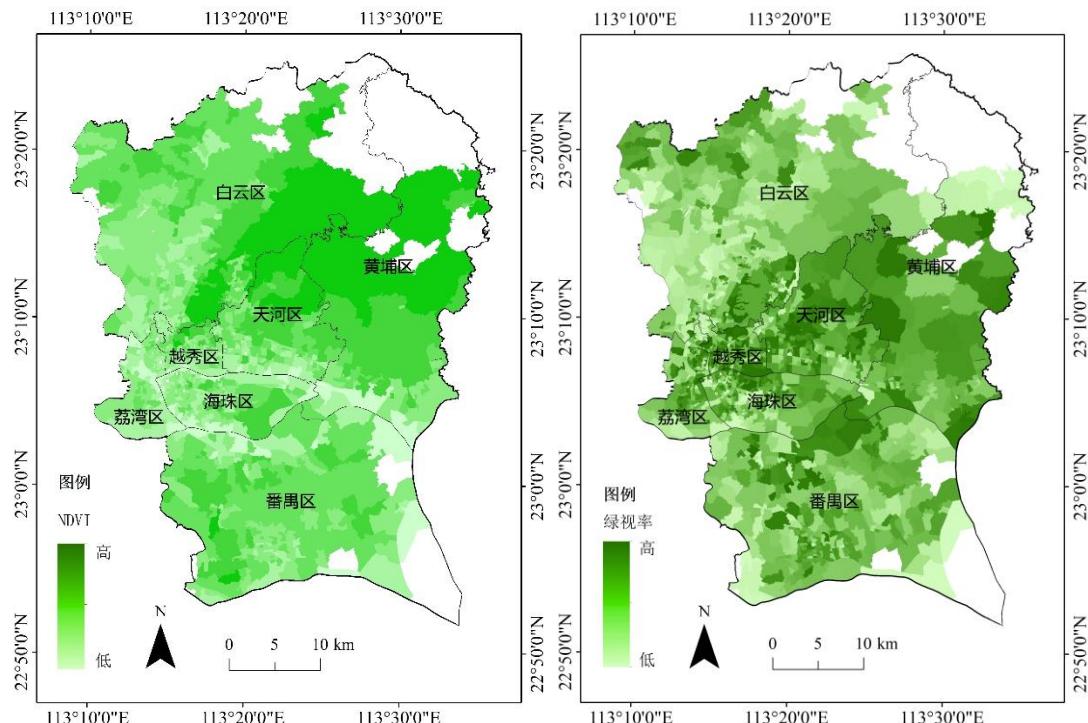


图 3-22 基于遥感影像的绿视率验证分析

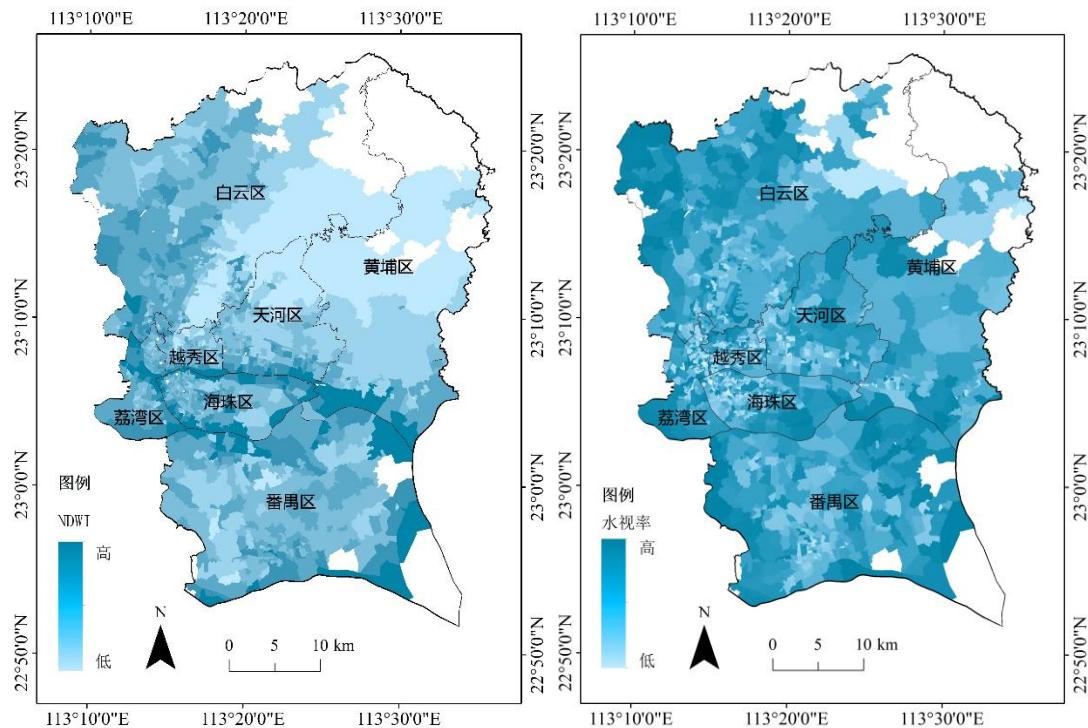


图 3-23 基于遥感影像的水视率验证分析

### 3.4 街景微环境空间集聚特征分析

根据 Tobler (1970)提出的地理学第一定律，相比于距离更远的对象，距离更近的对象的相关程度更强。理论上，每个个体，与该个体周围的个体之间存在一些共同特征，而且这些个体之间会有潜在的相互依赖性。这种没有明显表现出来的空间上的相互依赖性，就被称为空间自相关。

以社区尺度的微环境要素数据做全局空间自相关分析。分别选取 1 介 Queen 和 1 介 Rock 空间权重矩阵建立空间权重矩阵，然后进行全局 Moran 值的计算。 (Random999 次)。计算结果通过 Z 值检验 (P 值为  $0.001 < 0.05$ )，0.001 指进行蒙特卡洛模拟 999 次和真实分布 1 次所有 Moran' I 中，真实分布的 Moran'I 在一千个中排名第一。

由表 3-6 可以看出，利用两种权重矩阵计算得到的全局空间自相关结果较为类似。所有这些要素值都具有显著的空间集聚特征。从空间集聚结果来看，天空开阔指数的空间集聚特征最明显，其次是富裕度和建成环境特征。这些要素既包括自然要素，也包括人文要素。集聚特征最不明显的是垃圾度和水视率。可能原因之一是这两种微环境要素在街景中分布的都比较少。

表 3-6 全局空间自相关 Moran's I 值

评价值	Moran's I	Rank	P-value	Z-value
1 介 Queen				
多样性	0.3756	4	.001	24.9954
富裕度	0.555	2	.001	36.505
美丽度	0.365	5	.001	23.919
建成环境	.525	3	.001	35.549
天空	0.693	1	.001	46.981
绿视率	0.347	6	.001	23.941
水视率	0.153	9	.001	11.322
路灯比例	0.226	7	.001	15.305
垃圾度	0.181	8	.001	12.695
1 介 Rock				
多样性	0.375	4	.001	24.787
富裕度	0.557	2	.001	37.501
美丽度	0.368	5	.001	24.218
建成环境	.527	3	.001	33.989
天空	0.696	1	.001	45.555
绿视率	0.35	6	.001	23.283
水视率	0.154	9	.001	10.748
路灯比例	0.228	7	.001	15.449
垃圾度	0.181	8	.001	12.544

### 3.5 小结

微环境要素评估一直是一个重要的话题。如何更加准确地刻画反映社区失序状态的社区微环境对于安全感知和公共健康领域的研究至关重要。利用深度学习算法，本章节大规模刻画了广州市各类别微环境失序要素的空间分布特征，最后利用遥感影像完成了初步验证。研究结果表明：

#### (1) 建立了基于环境失序理论的街景微环境失序指标体系

基于环境失序理论，利用大规模街景开放数据和三种深度学习算法，本研究构建了一套街景微环境失序指标体系。三种街景图像挖掘的算法分别是获取感知失序环境要素的 SE-Inception 卷积神经网络模型，获取客观失序环境要素的全卷积网络 FCN-8s 图像分割模型和 Faster-R-CNN 目标检测模型。依据这一套微环境失序指标体系，可以分析城市研究相关问题，如城市犯罪分析、公共健康分析、城市贫困分析等。

#### (2) 刻画了广州市研究区微环境失序要素的空间分布特征

通过大规模的图像微环境得分评价，本章节得到了研究区微环境要素及微环境失序要素的空间分布特征，并基于遥感影像进行了初步验证。感知微环境方面，

以越秀和荔湾为代表的老城区，以及各大高校的集中区域、开放空间区域等是美丽、富裕程度较高的区域。而中心城区外围区域的美丽度和富裕度较低。这与现实的印象相符合。客观微环境方面，广州市天空比例较高的地区多位于城市郊区，而绿视率和建筑物比例较高的地方主要位于中心城区。这些客观微环境要素的空间分布也与认知相符合。整体上看，研究区建筑垃圾的分布很少，有垃圾图片占总图片的比例为 6.75%，说明广州市中心城区的清洁度水平较高。汇集到社区尺度，建筑垃圾分布的高值区域集中在番禺区中部和荔湾区南部。

#### （3）分析了广州市研究区微环境失序要素的空间统计特征

利用相关分析和空间自相关分析手段，本研究分析了研究区微环境要素及微环境失序要素的空间统计特征。研究发现，富裕度和美丽度、绿视率等要素的相关性较高，建筑物比例和天空比例、美丽度的相关性较高。建筑物比例越高，天空比例越小，街景美丽度得分也越小。其他微环境要素之间的相关性程度相对较低。建筑物比例和天空开阔程度的集聚特征明显，且呈负相关状态。而垃圾度和水视率集聚特征最不明显。

本章的研究成果可作为安全感知影响因素分析和空间模拟的重要输入变量，进而帮助城市规划部门检测城市区域安全感的动态变化，并评估随之带来的影响。同时，本章节研究说明，可以通过更新街景图像来监测城市微环境以及城市环境特征的动态变化。

本研究还有一定的改进空间。计算机视觉是当今热门的研究领域，各种深度学习算法层出不穷。比如，基于图像分割的方法，已经有分割效果更优的 PSPNet 金字塔池化模型，目标检测也有检测速度更快的 YOLO V3 网络。在以后的研究中，根据本研究的几种深度学习任务，在原有方法基础上进行改进，可以获取更加精准的城市失序微环境数据，以建立更科学的城市失序微环境评价体系。因此，在以后的研究中，如何有效获取更为精准的相关城市失序微环境指标，是需要攻克的新问题。



## 第四章 微环境要素对安全感知的直接作用分析

### 4.1 引言

一般认为，安全感知受个体因素和环境因素的共同作用（Wyant, 2008; Jing et al., 2020）。国内外一系列研究探讨了个体因素（性别、年龄、收入等）和环境因素（人口密度、犯罪率、环境失序特征等）对安全感知可能的影响。然而，客观环境特别是微环境因素对安全感知的影响评估的相关研究还不深入。主要可能原因有两点。其一，安全感知的研究大多数是犯罪学、社会学、心理学研究工作者参与的，这些学科领域主要关注于个体因素和社会环境因素。其二，技术手段上的限制导致了这方面的研究较少。一般情况下，安全感知的研究都运用了大批量的问卷，受访者来自于随机选取的多样化社区。以往要获取多个社区中诸如美丽度、绿视率、天空开阔程度、建筑物比例等环境要素较为困难，也难以通过大批量处理街景而得到多个社区的微环境指标。

同时，根据破窗理论(Wilson & Kelling, 1982)，环境失序因素是安全感知的重要影响指标。然而，环境失序要素的度量是一个还没有完全解决的问题。之前研究主要通过问卷调查的方式衡量感知的建成环境或社会环境失序，基于街景图像等客观数据源的客观环境失序指标的相关研究还很少(Mooney et al, 2017)。目前，虽然在公共卫生领域已经有研究检验客观环境失序与健康指标的关系，但是很少研究分析客观环境失序对安全感知的作用。

此外，一方面，作为中国背景下的特殊社会环境，户口因素区分了本地户口和外地户口，导致了个体在一系列社会福利方面的不平等 (Chan & Zhang, 1999)，以及导致了个体居住环境以及居住环境感知的差异 (Wu, 2012)，从而可能导致安全感知的差异。另一方面，当前社会环境和安全感关系的研究主要是基于西方社会背景开展的，这些研究大多基于被害模型、社会失序模型和社会融入模型等理论模型开展。然而，这些模型是否也适用于中国的社会背景下，还需要一系列以中国为案例地的研究去探索。

本章主要有四个研究目的。第一，根据上一章节计算出的基于街景的各种微环境指标，分析这些微环境指标与居民社区安全感知的关系。在新技术的支撑下，这不仅可以弥补环境和安全感关系理论研究上的不足，还可以作用于实际的城市

规划和城市管理工作。在分析基于街景的微环境与安全感知关系的同时，本研究分析街景环境要素的加入，是否会增强个体安全感知模型的解释能力，从而优化传统的个体安全感知模型。根据优化后的安全感知模型，有利于更准确地分析出影响居民安全感知的因素，从而在实践中通过环境设计预防犯罪来提升居民安全感。

第二，本研究尝试分析哪一种指标是最能反应社区环境失序程度的指标。社区环境有序程度和失序程度一直被认为是预测安全感知的关键要素。通过第四章的微环境要素提取，获取了大量微环境要素。其中，诸如垃圾度、美丽度、富裕度、绿视率等微环境要素在理论上都与社区失序程度有关，是可能的社区失序程度的指标。因此，本研究尝试分别加入这几种社区失序程度的可能指标，通过模型拟合优度的对比，选取最优的社区失序程度指标，以此完善模型。

除了以上主要目的，本研究还考虑中国背景下一些社会环境因素与安全感知的关系。比如，本研究会探索户口等中国特色的社会环境因素对安全感知的影响，以及西方安全感知分析模型在中国背景下的可适用性。

本章节的研究流程图如图 4-1 所示。

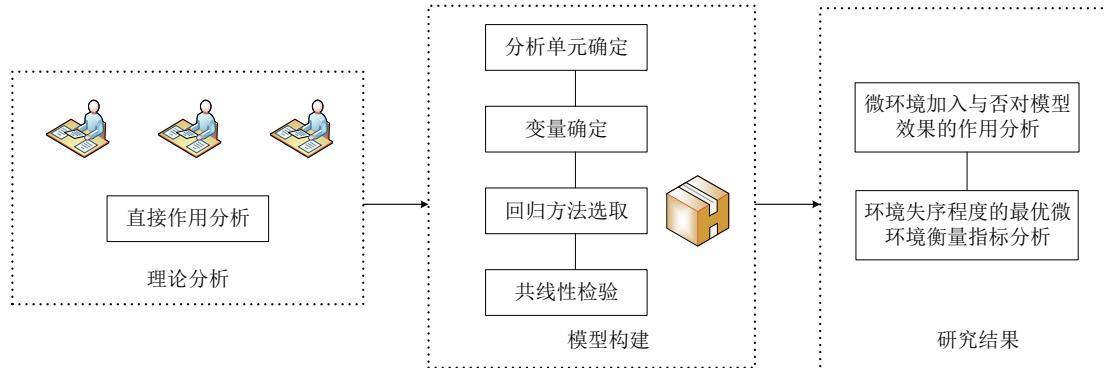


图 4-1 微环境要素直接作用分析流程图

## 4.2 研究方法

### 4.2.1 OLS 多元线性回归模型

#### 4.2.1.1 理论模型

多元线性回归是一种根据多个输入指标 X 解释结果指标 Y 的方法，它假设

多个输入指标  $X$  与  $Y$  之间有线性的关系。这种关系的数学表达式为:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \mu \quad (4-1)$$

式 4-1 中,  $Y$  是因变量,  $X_1, \dots, X_k$  是自变量,  $\beta_0$  是常数项,  $\beta_1, \dots, \beta_k$  是回归系数,  $\mu$  是误差项。

对于多元线性回归模型的参数估计, 在误差平方和为最小的基础上, 使用最小二乘法(Ordinary Least Square, OLS) 估算每一个回归系数, 以完成回归模型的建立。普通最小二乘法最为常用, 利用矩阵运算解得系数矩阵。计算原理为:

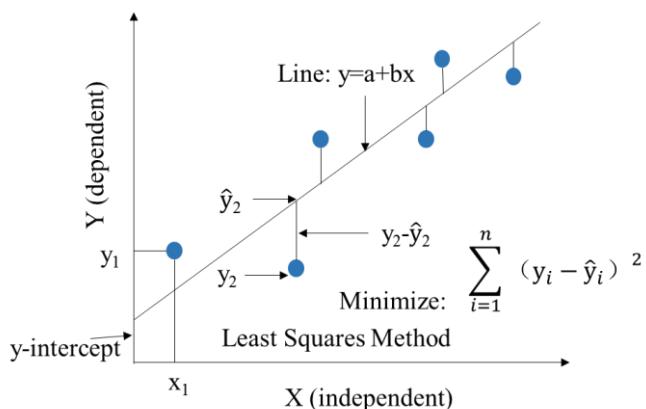


图 4-2 最小二乘法计算原理

#### 4.2.1.2 模型检验

确定模型后, 首先需要进行一系列模型检验以检测模型的正确性。第一个重要的检验指标是决定系数  $R^2$ , 它表示方程解释的变动在结果变量的总变异中所占的百分比。它的值越大, 回归方程对采样点数据的拟合越强, 所有自变量和因变量的关系越紧密。具体公式为:

$$R^2 = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \bar{y})^2} = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (4-2)$$

式中,

$$\begin{aligned} \sum (y - \hat{y})^2 &= \sum y^2 - (b_0 \sum y + b_1 \sum x_1 y + b_2 \sum x_2 y + \cdots + b_k \sum x_k y) \\ \sum (y - \bar{y})^2 &= \sum y^2 - \frac{1}{n} (\sum y)^2 \end{aligned} \quad (4-3)$$

第二个重要指标是估计的标准误差。它是指结果变量  $y$  的真实值和线性回归模型求出的预测值  $\hat{y}$  比较得到的误差值。较小的估计的标准误差表示方程较好的拟合效果。计算公式为:

$$S_y = \sqrt{\frac{\sum (y - \hat{y})^2}{n - k - 1}}$$

$$V_k = \frac{S_y}{y} \quad (4-4)$$

式里，k 表示回归方程中输入变量的个数。

第三个重要的检验指标是 F 检验。使用假设检验法，计算 F 统计量，可以检测出结果变量和输入变量之间是否存在统计意义的关系，如式 (4-5)：

$$F = \frac{\frac{\sum (\hat{y} - \bar{y})^2}{k}}{\frac{\sum (y - \hat{y})^2}{n - k - 1}} = \frac{\frac{R^2}{k}}{\frac{1 - R^2}{n - k - 1}} \quad (4-5)$$

按照显著性值 a，自由度 (k, n-k-1)，查询 F 分布表，可获得临界值  $F_a$ ，如果  $F > F_a$ ，表明回归有显著的效果；否则，回归方程没有显著的效果。

F 检验可以确定方程整体上是否达到显著性检验的要求。而要分析每个输入变量与结果变量之间的显著性关联，需要使用 t 统计量，

$$t_i = \frac{b_i}{S_y \sqrt{C_{ij}}} = \frac{b_i}{S_{bi}} \quad (4-6)$$

式中， $C_{ij}$  表示回归模型里，系数矩阵的逆矩阵  $(x'x)^{-1}$  对角线的第 j 个变量。

得到统计量  $t_i$  后，根据 a 值和自由度值，查询得到分区值  $t_a$  或  $t_{a/2}$ 。当  $t > t_a$  或  $t_{a/2}$ ，则回归系数  $b_i$  与 0 有显著差异；否则，与 0 没有显著差异。这样，通过确定接受原假设与否，从而判断某个自变量是否通过了显著检验。

#### 4.2.2 模型评价指标

对于回归方程的评估，有若干种拟合优度值。对于多元线性回归模型， $R^2$  可以用来判断模型的质量。在此基础上，本研究使用赤池信息量准则（AIC）和贝叶斯信息量准则（BIC）等统计指标来进一步综合判断模型的质量。

赤池信息量（Akaike information criterion, AIC）表示预测误差估计值，它将一组统计方程的结果相互比较，并将它们从最佳到最差进行排名（Akaike, 1987）。“最佳”模型将是既不欠缺也不过度的模型。AIC 的计算公式为：

$$AIC = 2k - 2\ln(\hat{L}) \quad (4-7)$$

其中，k 为模型中估计参数的数量。 $\hat{L}$  为模型似然函数的最大值。对于给定数据的候选方程集，AIC 最小的方程是首选方程。

贝叶斯信息量准则（Bayesian information criterion, BIC）也是用于在有限的

一组模型中进行模型选择的标准 (Schwarz, 1978)。BIC 最低的模型是首选。它部分基于似然函数，并且与 AIC 密切相关。BIC 的计算公式为：

$$BIC = k\ln(n) - 2\ln(\hat{L}) \quad (4-8)$$

其中，各参数含义与 AIC 的公式一致，n 表示观察数据的样本数。BIC 大于 AIC 的惩罚项。

### 4.2.3 变量设定

#### 4.2.3.1 分析单元

本章节研究的调研数据来源于 PPSGN 调查问卷中位于研究区的 82 个社区，这些社区分别属于 67 个街道。样本受访者总数为 1741 个。因此，从样本分布来看，基于空间分布上的代表性，又有人口统计结构上的代表性。样本的社区如下图 4-3 所示。

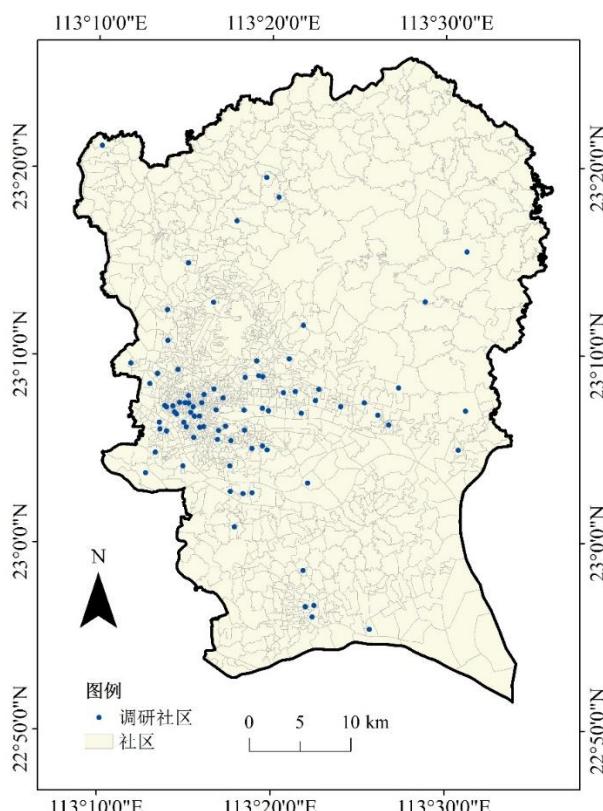


图 4-3 样本社区分布

#### 4.2.3.2 因变量

与之前国内外安全感知研究类似 (Wyant, 2008; Hanslmaier, 2013)，本研究的

因变量“居民安全感知”由 5 个指标构成，通过李克特五点量表的形式调查受访者对遭遇一系列被害场景的害怕程度。相比于单一题项，复合题项的设置，有助于从多维度确保受访者安全感知测量的准确性。由于问卷题项值越高代表担忧遭遇受害的水平越高，所以以“不安全感”为具体因变量名称。具体问题如下：

关于您对社区/小区及出行安全的评价，您是否同意以下观点？

No.	问题	一点都不 不担心	有点担心	一般	比较担心	很担心
1	不在家时，担心家里被盗(入室盗窃)	(1)	(2)	(3)	(4)	(5)
2	担心我的随身财物在公共场所被偷	(1)	(2)	(3)	(4)	(5)
3	晚上 10 点后在社区/小区散步会感到 害怕(担心被攻击或者被陌生人骚扰)	(1)	(2)	(3)	(4)	(5)
4	我害怕被抢劫	(1)	(2)	(3)	(4)	(5)
5	我害怕被骚扰	(1)	(2)	(3)	(4)	(5)

每一个受访者的不安全感由这五项指标的均值构成。信度分析显示出该指标体系具有非常强的一致性（信度分析指标  $\alpha=0.91$ ）。为了验证居民安全感知是否符合正态分布，本研究对因变量进行了峰度和偏度分析。结果显示，峰度和偏度都小于 1（偏度值为 0.265，SE=.059；峰度值为-0.348，SE=.117），因此，基于最小二乘法的多元线性回归模型适用于此类数据的回归分析，故本研究拟采取 OLS 多元线性回归模型进行回归检验<sup>5</sup>。

#### 4.2.3.3 个体属性数据

影响居民安全感知的因素可归纳为个体因素和环境因素。个体因素是如居民的年龄、性别等个体特征。以下对个体因素进行简单描述。

##### （1）身体脆弱性指标

性别（分类变量，女性记为 1，男性记为 0）。年龄（定序变量，从 1 到 4。18 至 30 岁记为 1，30（含）至 45 岁记为 2，45（含）至 60 岁记为 3，60 岁（含）及以上记为 4）。

##### （2）社会脆弱性指标

<sup>5</sup> 一个更理想的方法是使用多水平回归模型进行分析，因为它可以考虑到个体层和社区层变量之间的变异。但是，在本文研究中，每个社区平均只有 20 个样本，作为社区层的样本数量来说可能不足。因此，本研究没有选择使用更复杂的多水平回归模型。

户口，分类变量，0 代表非本地户口，1 代表本地户口。个体收入（定序变量，从 1 到 7。1 表示 1000 元以下；2 表示 1000-1999 元；3 表示 2000-2999 元；4 表示 3000-4999 元；5 表示 5000-7999 元；6 表示 8000-9999 元；7 表示 10000 元以上），测量了个体的富裕程度。此外，测量了个体受教育水平（定序变量，从 1 到 7。受教育程度随着取值的增加而增加）。

### （3）受害模型

与之前文献一致(Brunton-Smith & Sturgis, 2011)，两种受害经历纳入本模型中的变量，分别是直接受害经历和间接受害经历。直接受害经历是询问受访者在过去三年内是否亲自遭遇过各种犯罪被害事件（1 记为有直接受害经历，0 记为无直接受害经历）。间接受害经历是询问受访者在过去三年内是否听说亲戚、朋友、邻居等有遭遇过各种犯罪被害事件（1 记为有间接受害经历，0 记为无间接受害经历）。

### （4）感知的邻里环境

① 感知的建成环境失序。邻里失序是安全感知的一个关键变量。感知的建成环境失序包括 4 项问题，以取均值的形式获得。问题是询问受访者对以下现象在社区中的严重程度的评价：1) 损坏的公共设施；2) 杂乱的垃圾和乱停放的汽车；3) 涂鸦和乱贴的广告；4) 吵闹的邻里环境。回答的选项从 1（一点儿也不严重）到 5（非常严重） $(\alpha=0.89)$ 。

② 感知的社会环境失序。相应地，感知的社会环境失序也包括 4 项问题，以均值的形式获得。问题是询问受访者对以下现象在邻里中的严重程度的评价：1) 青少年聚集在街道上闲逛；2) 邻居之间容易冲突；3) 可疑的陌生人在周边活动；4) 晚上醉酒的人在街上闲逛。回答的选项从 1（一点儿也不严重）到 5（非常严重） $(\alpha=0.85)$ 。

③ 感知的非正式社会融入。非正式社会融入与社区集体效能密切相关，通常被认为是与抑制不安全感有关。它包括四项，通过以下问题测量：请您评估下以下现象在小区周围出现的频率：1) 邻居之间见面互相打招呼；2) 邻里之间互相访问；3) 邻里之间互相借东西；4) 邻居之间的关系很融洽。越高的值代表非正式社会融入程度越高 $(\alpha=0.84)$ 。

④ 非正式的社会控制。与之前文献一致(Yuan & McNeely, 2017)，本研究中的非正式社会控制的问题如下：1) 如果有小朋友在小区附近乱涂乱画，小区邻

居会试图阻止吗？2) 如果在你家住处外面有人打架或者有人被打，小区邻居会介入阻止吗？3) 如果有陌生人在你小区附近闲逛，小区邻居可能会有什么举动吗？4) 如果有认识的小朋友逃学，小区邻居如果发现了，会告诉小朋友的家人吗？非正式社会控制的值由这四项值的平均值构成(值的范围在1(低社会控制)到5(高社会控制)波动)，内在一致性程度较高 ( $\alpha=0.84$ )。

⑤ 感知的社区警力。增加的社区警力被认为与较低的不安全感有关。感知的社区警力一般通过多个指标来测量，但是由于问卷所设题项限制，本研究使用一个单独的维度来测量感知的社区警力。问题项为：您在小区或小区附近经常看到警察或者保安巡逻吗？值越高，代表感知的社区警力越强。

### 4.2.3.4 环境属性数据

本研究中的环境属性数据原计划包括基于统计数据的人口普查指标、基于遥感数据的自然环境指标、基于街景数据的街景环境指标、基于POI数据的设施点环境指标共15项指标。然后，将这些环境指标赋予到每一个被试者数据属性上。

#### (1) 社区人口普查指标

① 人口密度。人口密度是计算社区总人口和社区土地面积的比例，单位为人/每平方千米。

② 受教育水平。受教育水平是计算社区受高等教育人口(大专及以上)占总人口的比例。

③ 外来人口比例。外来人口比例是计算社区外来人口(非本市户籍人口)占总人口的比例。

④ 犯罪率。犯罪率是计算2014年和2015年两年期间社区犯罪案件数量(“两抢两盗”，即抢劫案件量和盗窃案件量)与总人口的比值，然后除以2。

以上数据源主要来自于2010六普统计数据以及本研究团队和广州市公安局合作获得的广州市犯罪数据。

#### (2) 基于遥感数据的自然环境指标

基于之前文献(Mouratidis, 2019)，社区的绿地水平可能与居民的安全感知有关。因此，本研究将反应空间尺度下植被稀疏程度的NDVI指数作为影响安全感的一个变量。NDVI指数是基于Landsat-8卫星影像数据，通过ENVI软件进行辐射定标、大气校正、融合、波段计算后获得栅格化的NDVI图层数据，再通过

ArcGIS 软件将栅格 NDVI 数据转换为基于社区的 NDVI 指数值，计算方式为均值法。其中，NDVI 的计算公式（Kaufman & Tanre, 1992）如下：

$$\text{NDVI} = \frac{\text{NIR}-\text{VIS}}{\text{NIR}+\text{VIS}} \quad (4-9)$$

健康的植被反射近红外光(near-infrared light, NIR)，但吸收可见光(VIS light)。NDVI是基于光谱的可见(红色)和近红外部分的地面反射率来指示绿色的植被，通过将 NIR 和可见辐射 (VIS) 的差除以 NIR 和可见辐射 (VIS) 的总和来计算得出。

### (3) 设施点指标

根据获取的广州市 POI 设施点数据，基于之前文献(张延吉, 秦波 & 唐杰, 2017)，本研究筛选出两个可能与安全感知密切相关的指标，分别是社区商业设施比例和社区政府机构设施比例。其计算方法是，特定人口普查社区内的该类别 POI 设施数目除以该社区面积(平方千米)。一般认为，政府机构设施比例越高，居民的不安全感越低；商业设施比例越高，居民的不安全感可能越高。

### (4) 街景微环境指标

街景微环境指标包括人口普查社区尺度的街景失序程度（美丽度、富裕度、绿视率指标之一）、街景天空开阔程度、街景路灯比例、街景垃圾度等指标。指标的具体含义已在上一章节描述，故不再赘述。

街景微环境指标的选择是在理论分析、VIF 共线性检验和相关系数检验上完成的。首先，由于街景照片中的水视率指标普遍较少或为 0，因此本研究没有将可能影响安全感知的街景水视率指标纳入模型。其次，街景微环境中的一系列指标相关性较高，包括多样性、建筑物比例、天空开阔程度、绿视率、美丽度、富裕度等。垃圾度、绿视率、美丽度、富裕度等四个指标作为社区失序程度的可能有效指示指标，将依次纳入模型，不存在相关性和共线性的问题。据此，根据上一章模型实验，由于建筑物比例与天空开阔程度相关度较高，多样性与绿视率、美丽度等指标相关度较高，本研究将建筑物比例指标、多样性指标删除。

#### 4.2.3.5 数据检验

本研究分别进行了相关分析和多重共线性检验以避免所有自变量中存在多重共线性问题。

### (1) 相关性分析

采用皮尔逊相关系数，实现相关分析并对结果进行了可视化（图 4-4）。结果显示，一些显著的正向影响、负向影响分别存在一些变量之间。所有变量之间（因美丽度、富裕度、绿视率 3 个指标之间的相关性高，此处只纳入绿视率作为客观失序指标进入相关分析）的相关性小于 0.5，变量都可以纳入回归模型进行检验。

### (2) 多重共线性分析

在此基础上，进行 OLS 多元线性回归分析前，本文使用 SPSS Statistics 22 对各变量进行了方差膨胀因子（VIF）分析。结果表明，各模型的变量的方差膨胀因子（VIF）值均小于 4。多重共线性问题不存在。

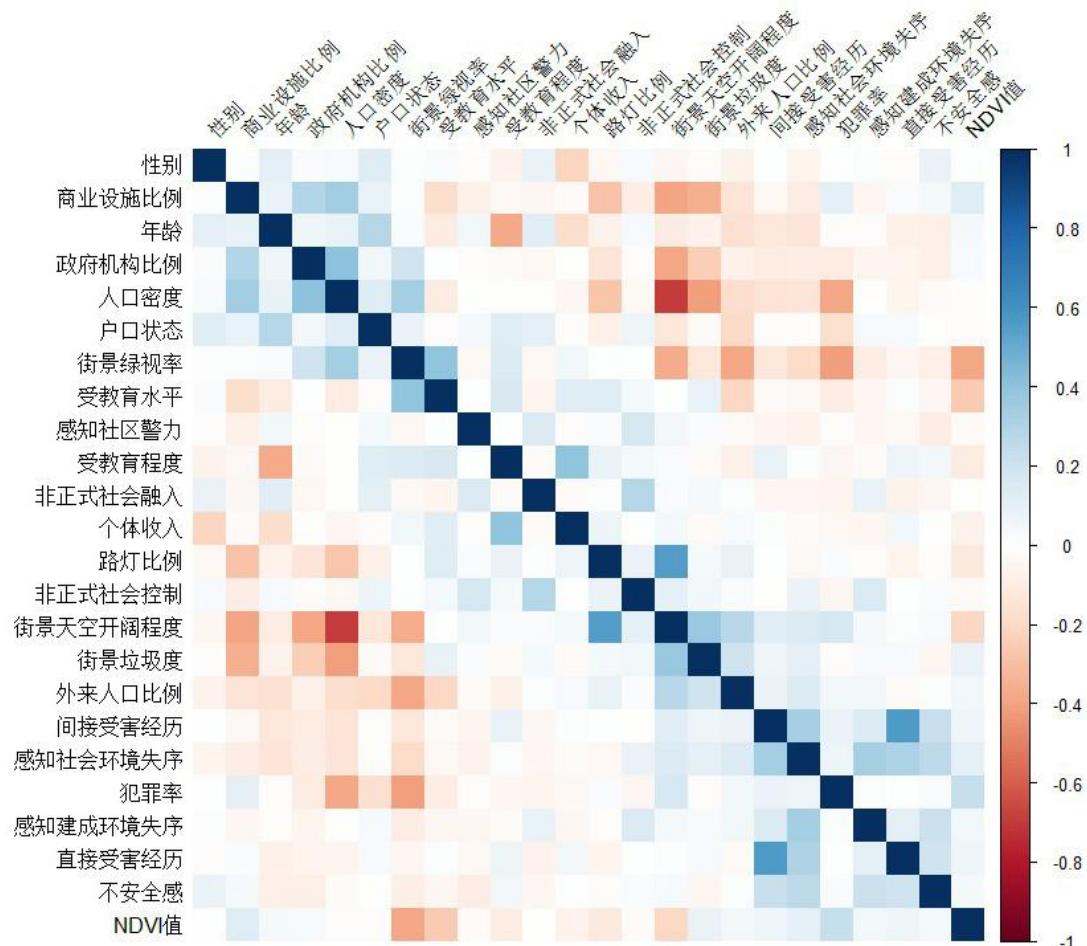


图 4-4 变量相关分析图

## 4.3 研究结果

### 4.3.1 描述统计结果

表 4-1 样本变量描述统计分析

变量	N	最小值	最大值	平均值	标准差
<b>个体因素</b>					
年龄	1741	1	4	2.04	0.984
性别	1741	0	1	0.47	0.499
户口	1741	0	1	0.48	0.500
教育程度	1741	1	7	3.93	1.174
个人收入	1741	1	7	3.55	1.378
直接受害经历	1741	0	1	0.38	0.485
间接受害经历	1741	0	1	0.45	0.498
感知的建成环境失序	1741	1	5	2.91	0.808
感知的社会环境失序	1741	1	5	2.06	0.817
非正式的社会融入	1741	1	5	3.42	0.793
非正式的社会控制	1741	1	5	3.02	0.782
感知的社区警力	1741	1	5	3.23	0.879
<b>环境因素</b>					
受教育水平	82	0.003	0.863	0.144	0.209
外来人口比例	82	0.000	0.900	0.303	0.262
人口密度	82	187.585	115992.113	31381.713	29529.829
商业设施比	82	0.211	1921.283	181.185	283.050
政府机构比	82	0.000	776.074	54.423	105.158
遥感 NDVI 值	82	-0.275	0.375	-0.035	0.103
犯罪率	82	102.56	8801.12	1594.94	1755.57
路灯比例	82	0.263	0.811	0.544	0.131
天空开阔程度	82	0.063	0.463	0.291	0.087
街景垃圾度	82	0.000	0.200	0.057	0.046
街景绿视率	82	0.068	0.444	0.238	0.072
街景美丽度	82	0.430	0.558	0.489	0.023
街景富裕度	82	0.472	0.606	0.522	0.025

本章节的变量描述统计如上表所示。可知，年龄的均值接近于等级 2，表示大部分人口的年龄区间在 30-45 岁。男性样本数要略多于女性。非广州户口的认

识略多于广州户口。这些人口统计特征显示，PPSGN 的调研结果与统计数据基本一致，调研问卷具有理想的数据质量。此外，有直接或间接受害经历的人群较少。受访者感知的建成或社会环境失序程度较低，非正式社会融入和控制，以及感知的社区警力水平较高。环境方面，研究区的人口密度较高，符合大城市的稠密人口分布状态。社区的街景垃圾度水平较低，街景绿视率水平的均值为 0.238，街景美丽度和富裕度的得分较为接近。

### 4.3.2 微环境对安全感知的直接作用分析

表 4-2 微环境对不安全感的直接作用分析

	Model 1		Model 2		Model 3		Model 4	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
常量	2.350**	0.126	1.740**	0.193	1.749**	0.204	2.153**	0.293
个体								
年龄	-0.067*	0.027	-0.036	0.027	-0.042	0.027	-0.041	0.027
性别	0.189**	0.048	0.209**	0.046	0.219**	0.046	0.218**	0.046
户口	-0.002	0.050	-0.015	0.049	-0.037	0.05	-0.04	0.05
教育程度	0.009	0.024	0.022	0.023	0.023	0.023	0.026	0.023
个人收入	0.000	0.019	0.008	0.018	0.013	0.018	0.012	0.018
直接受害经历	0.232**	0.057	0.156**	0.056	0.146*	0.056	0.148**	0.057
间接受害经历	0.300**	0.056	0.189**	0.055	0.184**	0.056	0.181**	0.056
感知建成环境失序			0.168**	0.030	0.164**	0.03	0.164**	0.03
感知社会环境失序			0.195**	0.031	0.205**	0.031	0.204**	0.032
非正式的社会融入			-0.053#	0.030	-0.056#	0.03	-0.056#	0.03
非正式的社会控制			0.011	0.030	0.023	0.03	0.025	0.031
感知的社区警力			-0.064*	0.026	-0.06*	0.026	-0.062*	0.026
社区结构								
社区教育水平					-0.226*	0.111	-0.196#	0.119
社区外来人口比例					-0.092	0.093	-0.136	0.099
社区人口密度					-0.000	0	-0.000	0
社区商业设施比					0.000**	0	0.000**	0
社区政府机构比					-0.001**	0	-0.001**	0
遥感 NDVI 值					0.337	0.237	0.689	0.298
社区犯罪率					-0.000	0	-0.000	0
社区微环境								
社区路灯比例							0.344	0.226
街景天空比例							-0.875	0.551
街道垃圾度							-0.618	0.542
街景绿视率							-1.007*	0.479
R <sup>2</sup>	0.071		0.13		0.143		0.147	
AIC	4789.833		4723.427		4716.015		4674.056	
BIC	4833.531		4788.973		4819.797		4788.763	

注： \*\* p<0.01, \* p<0.05, # p < 0.1; Coeff 为非标准化系数。

运用 StataSE 15，选取最小二乘法进行参数估计，完成多个 OLS 回归模型的计算。在表 4-2 中，对比 Model 4 与其他模型，Model 4 的 R 方最大 (0.147)。随着感知环境、社区环境、微环境因素的逐步加入，模型的 R 方不断增加，表明模型的拟合优度不断增强（社会科学相关解释性回归分析中的 R 方普遍不高，安全感知线性回归分析中，R 方多介于 0.03~0.3。因此，相比之前文献 (Wyant, 2008; Scarborough et al., 2010; Hanslmaier, 2013)，本研究各模型的 R 方处于正常偏高的区间，模型效果良好）。根据 AIC 和 BIC 的值也可以看出，在加入微环境因素后，Model 4 的 AIC 值和 BIC 值相较于其他模型，都有减少，说明加入微环境因素后回归模型的解释度更好。此外，一些传统指标显著影响安全感知。女性比男性更容易有不安全感，直接和间接受害都显著增强了不安全感，感知环境要素如感知建成和社会环境、非正式社会融入、社区警力都与安全感知显著相关。社区因素方面，社区教育水平越高，不安全感水平越低。社区商业设施比是促进居民不安全感的因素，而社区政府机构比是促进居民安全感的因素。

### 4.3.3 环境失序程度的最佳微环境衡量指标

表 4-3 显示了 4 个模型的分析结果，以此检验用于环境失序程度的最佳微环境衡量指标。Model 5、6、7、8 分别加入了四个可能反应街景失序程度的指标到回归模型中。结果显示，模型 6 中，街景绿视率是显著影响安全感知的指标（系数 Coeff=-1.091, P<0.05）。即，随着社区街景绿视率水平的增加，社区居民的不安全感会减弱。然而，模型 7 显示，虽然街景美丽度的增加会减弱居民的不安全感，但是这种关系没有统计意义上的显著性。街景富裕度和垃圾度也没有统计意义上显著地影响居民不安全感。对比这四个模型的拟合优度指标，发现模型 6 的 R 方值最高，AIC 值和 BIC 值最低，表明模型 6 的模型解释效果最好。因此，模型显示，街景绿视率是反映街景环境失序程度的最优指标。

表 4-3 环境失序程度的最佳微环境衡量指标分析

	Model 5		Model 6		Model 7		Model 8	
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
常量	1.822**	0.247	2.133**	0.292	2.167**	0.751	1.111	0.89
个体								
年龄	-0.044	0.027	-0.040	0.027	-0.043	0.027	-0.045	0.027
性别	0.219**	0.046	0.218**	0.046	0.219**	0.046	0.219**	0.046
户口	-0.032	0.050	-0.044	0.050	-0.033	0.05	-0.031	0.05
教育程度	0.022	0.023	0.025	0.023	0.022	0.023	0.021	0.023
个人收入	0.013	0.018	0.012	0.018	0.013	0.018	0.012	0.018
直接受害经历	0.152**	0.057	0.145**	0.056	0.151**	0.057	0.153**	0.057
间接受害经历	0.178**	0.056	0.186**	0.056	0.179**	0.056	0.178**	0.056
感知建成环境失序	0.166**	0.030	0.163**	0.030	0.166**	0.03	0.166**	0.03
感知社会环境失序	0.206**	0.032	0.206**	0.031	0.205**	0.032	0.208**	0.032
非正式的社会融入	-0.053#	0.030	-0.059*	0.030	-0.053#	0.03	-0.052#	0.03
非正式的社会控制	0.021	0.031	0.028	0.031	0.022	0.031	0.02	0.031
感知的社区警力	-0.060*	0.026	-0.061*	0.026	-0.061*	0.026	-0.059*	0.026
社区结构								
社区教育水平	-0.262*	0.115	-0.184	0.119	-0.25*	0.118	-0.272*	0.116
社区外来人口比例	-0.076	0.095	-0.143	0.099	-0.088	0.098	-0.057	0.098
社区人口密度	0.000	0.000	0.000	0.000	-0.000	0	-0.000	0
社区商业设施比	0.000**	0.000	0.000**	0.000	0.000**	0	0.000**	0
社区政府机构比	-0.001**	0.000	-0.001**	0.000	-0.001**	0	-0.001**	0
遥感 NDVI 值	0.397	0.264	0.736*	0.295	0.476#	0.31	0.282	0.298
社区犯罪率	0.000	0.000	0.000	0.000	-0.000	0	-0.000	0
社区微环境								
社区路灯比例	0.205	0.216	0.374#	0.225	0.237	0.226	0.133	0.233
街景天空比例	-0.363	0.495	-0.971#	0.545	-0.44	0.52	-0.075	0.604
环境失序指标	垃圾度	-0.794	0.536		-1.091*	0.473		
	绿视率						-1.023	1.386
	美丽度							.740
	富裕度							1.490
R <sup>2</sup>	0.144		0.146		0.144		0.143	
AIC	4677.43		4674.262		4679.099		4679.408	
BIC	4803.506		4799.893		4804.73		4805.039	

注: \*\* p<0.01, \* p<0.05, # p < 0.1; Coeff 为非标准化系数。

## 4.4 小结

根据一系列安全感知理论和环境犯罪学理论, 本章节分析了广州市居民安全感知的影响因素, 并比较了衡量环境失序程度的几种指标。得到如下结论:

(1) 微环境因素的加入有利于增强安全感知模型的解释效果

本研究通过比较加入微环境和未加入微环境的模型拟合效果, 发现加入微环境的模型评价指标 (R<sup>2</sup>, AIC 和 BIC 值) 要优于未加入微环境的相关评价指标。该发现证明了街景微环境对于居民安全感知的重要作用, 有助于今后完善安全感

知模型的构建，以检验出更真实的安全感知影响因素。

### (2) 街景绿视率是安全感知模型中反映环境失序程度的显著指标

通过比较街景绿视率、美丽度、富裕度、垃圾度等四种反映环境失序程度的指标，研究发现街景绿视率是显著影响不安全感的指标，而且基于街景绿视率的不安全感知模型的拟合效果最优。因此，本研究证实了街景绿视率是反映环境失序程度的有效指标。这一发现进一步丰富了环境和安全感关系的理论研究。同时，本研究的这些发现也进一步拓宽了安全感知的研究领域。在接下来的研究中，厘清客观环境对安全感知的影响机制是一个有意义的研究领域。

### (3) 一系列基于西方理论模型的安全感知影响因素被验证

根据安全感知的经典理论模型，本研究验证了脆弱性模型、受害模型、失序模型、社区警力模型等在中国安全感知研究中的适用性。研究表明，女性、受害经历、感知的建成失序或社会失序程度高与居民增加的不安全感显著相关。感知的社区警力与居民减弱的不安全感显著相关。一系列社区因素中，与简雅各布斯的街道眼理论(Jacobs, 1992)和 Newman (1972)的防卫空间理论一致，本研究验证了街道绿视率与减弱的居民不安全感之间显著的关系。这一发现也与 Maas et al. (2009)的发现一致。社区的街景绿视率越高，表示该社区的自然环境较好，社区有序程度高。此外，居民生活在街景绿视率高的区域，会缓解居民的负面情绪，减弱居民的暴力情绪，提升居民的幸福感。最后，街景绿视率高的区域，为居民活动交流提供了便利，可能导致社区居民的社会融入程度增加，进而导致居民的不安全感下降。

但是，与街道眼理论相反，本研究发现社区商业设施比例是促进不安全感的显著指标。这可能是由于在社区商业占比高的区域，人流量大，居民遭遇被害(被偷、被抢)的风险就越大，从而导致个体的安全感降低 (Wilcox, Quisenberry & Jones, 2003)。社区的政府机构比例越高，居民的不安全感知越低。这也与常识相符合，同时与张延吉,秦波 & 唐杰 (2017) 的发现一致。政府机构占比越高的区域，社会秩序越有序，客观治安条件越好，社区居民的安全感水平越高。此外，根据社会生态学理论，本研究验证了社区一些结构指标(如社区受教育水平和社区政府机构比例)与安全感知之间的关联。社区受教育程度越高，该社区居民的不安全感越低。社区受教育程度高的区域，通常是社会失序程度较低的社区 (Jing et al., 2020)，因此可能导致了较低的风险感知水平。

(4) 影响安全感知的主要因素与之前文献一致，但是仍存在一些不同

一些影响因素没有得到验证。年龄、受教育程度、收入等身体脆弱性指标和社会经济脆弱性指标并没有统计意义上显著地与安全感知相关，不过，这也与一些文献的结果类似 (Liu et al., 2009; Wyant, 2008)。社会融入模型中的非正式社会控制也没有统计意义上显著地影响安全感知。这也说明，虽然安全感知的理论模型是适用于不同文化和社会背景下的研究地，但是具体不安全感的影响因素（包括环境因素）是会因社会文化背景而异的。与一些文献类似 (Scarborough et al., 2010)，本研究没有发现犯罪率与居民安全感知统计意义上的显著关系。这可能是由于犯罪活动发生在特定时间和特定地点，大多数居民没有直接遭遇犯罪，所以对于实际的犯罪感知不强烈。此外，其他微环境因素和一些社会结构指标（外来人口比例和人口密度等）没有显著地影响居民的不安全感。这些因素可能存在间接影响作用，还需要进一步探讨。

在实际的城市治安防控工作中，可以通过本章节的研究结论作用于城市治理。比如，通过提升街景绿视率的方法提升街区环境的风貌，进而提升居民的社区安全感知。就街景绿化项目来看，本研究的结论支持在广州城市区域种植街景植被可以增强居民的安全感知。沿街的树木和草地具有较高的可视率，可以广泛地被种植。维护良好的植被通常被认为与居民降低的被害恐惧感有关，所以种植的植被还需要加强日常维护。由于街景指标还可以作为一种领地标志 (sign of territory)，在主要道路节点种植植被可能带来更多益处。最后，高密度、高混合化的城市社区可以让居民感到更安全。因此，在城市规划和城市管理中，可以在一定程度上增加高密度、高混合化的社区，让居民能够拥有更高的安全感和归属感。

## 第五章 微环境要素对安全感知的间接作用分析

### 5.1 引言

现有的大多数安全感知文献是分别研究客观环境和主观环境对于居民安全感知的影响，发现了一系列与安全感知有关的主客观环境因素。然而，除少量研究外，该领域还缺乏耦合客观环境和主观环境对于安全影响因素分析的研究。因此，在上一章节研究客观微环境因素对安全感知的直接效应基础上，本研究进一步探讨客观微环境因素对安全感知的间接效应和总效应。本章节拟构建“客观环境-主观环境-安全感知”的影响机制，采用中介效应、调节效应、交互效应分析方法，对该机制进行实验分析，以探讨相关街景微环境的直接效应和间接效应机制。

本章节研究的核心目的是通过分析街景微环境的间接作用，为安全感知预测的输入变量选择提供理论支撑，也为安全感知的预测方法选择提供理论指导。通过中介效应分析，探讨出客观微环境是如何通过感知环境影响安全感知，可以从理论上解释选择客观微环境加入安全感知预测模型的原因。通过调节效应分析，探讨一些客观微环境对安全感知是否虽然没有直接作用，但是和其他因素共同作用于安全感知（交互效应或调节效应），可能存在间接作用，从而从理论上解释之后个体安全感预测时加入一些不显著的微环境的原因。综合开展中介效应、调节效应、交互效应、有调节的中介效应的分析，可以探测出安全感知是一个由多种因素相互作用的复杂机制，从而可以指导安全感知的预测。传统的线性回归模型的预测效果可能不够好，而需要选取其他可以处理复杂内在联系机制的非线性模型。

以下是具体的研究目的。

第一个目的，探讨街景微环境对居民安全感知的影响是否受主观环境的影响（中介作用）。由上一章节可知，环境因素对安全感知的作用区分为客观环境影响和感知环境影响两个方面。Brunton-Smith, Jackson & Sutherland (2014)的研究表明，客观环境（如社会经济水平、种族异质性等）会影响居民的集体效能感知和失序程度感知，从而影响居民的安全感知程度。也就是，一些客观环境对主观安全感知的影响，有一部分的作用机制是通过其它感知的环境指标影响的。以街

景绿视率为例，一些研究显示，绿视率可能是社区失序和集体效能的重要指标。同时，邻里失序和集体效能是影响居民安全感知的关键指标。因此，街景绿视率可能影响了邻里失序和集体效能，进而影响了居民安全感知。本研究拟选定特定街景微环境指标，探讨特定街景微环境指标对安全感知的影响是否受感知的集体效能和社区失序因素中介。

第二个目的，探讨街景微环境对居民安全感知的影响是否受主观个体因素的调节（调节作用和交互作用）。一系列研究表明，客观环境要素，如犯罪率、社区种族异质性等，会与居民的个体属性如性别、年龄、受害经历等产生调节作用或调节作用。比如，同样生活在环境失序高的社区，非白人相比于白人可能有更高的不安全感，女性相对于男性也更可能有更高的不安全感（Brunton-Smith & Sturgis, 2011; Snedker, 2015）。此外，有研究（Hunter & Baumer, 1982）也表明，生活在社会融入程度高的社区，周围街道人流量与居民的安全感知没有显著的关系；生活在社会融入程度低的社区，周围街道人流量越大，居民的不安全感显著增加。因此，客观环境和个体因素及个体感知因素可能存在调节作用或交互作用。本研究拟探讨特定街景微环境指标对安全感知的影响是否受个体因素调节，以及特定微环境指标和个体因素是否有共同交互作用以影响居民安全感。

第三个目的，探讨特定街景微环境因素对居民安全感知的影响是否存在有调节的中介作用。心理学的研究发现，除了单一的中介效应和调节效应，一些因素之间同时存在着这两种影响机制，即有调节的中介作用或有中介的调节作用。比如，林爱华，乔红芳 & 沈利生（2020）发现，政府环境规制对绿色发展水平有直接效应和间接效应。环境规制是否能够促进绿色发展水平，取决于环境规制的哪种形式成为主导。在涉及客观环境-主观环境-安全感知的研究中，还没有研究同时涉及到调节效应和中介效应的交互。进一步同时考虑调节和中介效应，可以揭示出客观环境-主观环境-安全感知之间更真实的关系和机制。

本章节以街景微环境中的绿视率和街道路灯比例作为客观微环境指标的代表，以感知的建成环境失序为感知环境指标，以直接受害经历和非正式社会融入为调节变量，以安全感知为因变量，综合分析客观环境-主观环境-安全感知之间的机制。

本章节的研究流程图如图 5-1 所示。

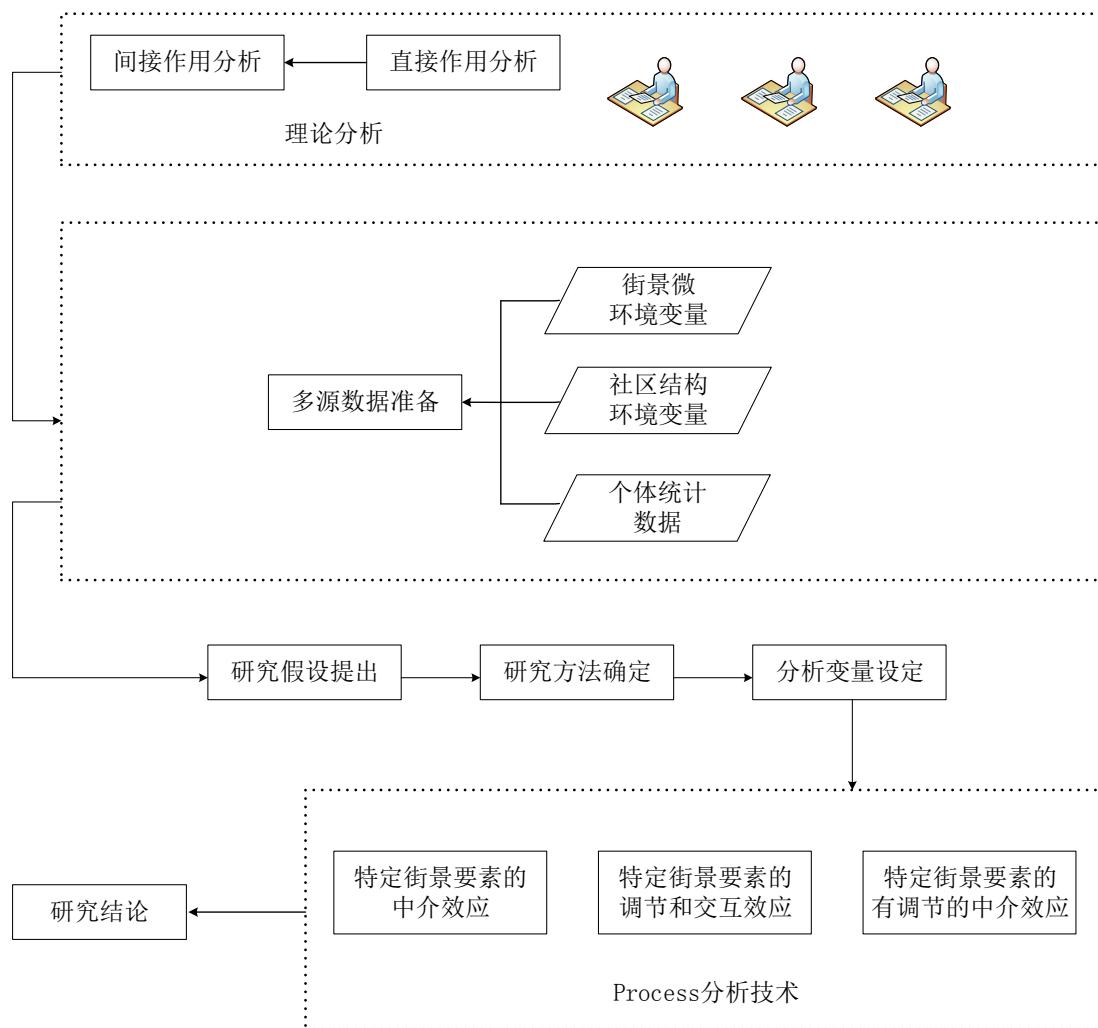


图 5-1 微环境要素间接作用分析流程图

## 5.2 研究假设

### 5.2.1 街景绿视率对安全感知的作用受建成环境失序因素中介

一系列研究已经表明了更多的居住区绿化可以减少社区犯罪率 (Maas et al., 2009; Mouratidis, 2019)。一些学者还发现植被与环境失序之间的关系 (Brunson, Kuo & Sullivan, 2001; Ellaway, Macintyre & Bonnefoy, 2005)。比如，在有许多维护良好植被的社区中，更少的失序问题被报告。同时，文献表明居住区植被对居民的幸福感和身体健康状态都有积极的影响 (Thompson et al., 2012; Richardson & Mitchell, 2010)。在安全感知影响因素的研究中，诸如失序、心理幸福感为代表的脆弱性特征都被认为是显著地影响了居民的安全感知。因此，本研究拟做出如

下假设：

H1：街景绿视率影响安全感知，同时这种影响会被感知的建成环境失序因素中介。

### 5.2.2 街景绿视率对安全感知的作用受直接受害经历因素调节

Maas et al. (2009) 关于居住区植被和社会安全感关系的研究中发现，居住区植被对居民社会安全感的影响受到所在区域城市化水平、受访者性别和年龄等因素的调节。比如，在高度城市化的区域，女性和老年人会有更高的社会安全感，但是男性的社会安全感较为稳定，不受城市化水平高低的影响。

直接受害经历，被发现会与一些因素交互，如自我适应(Rühs, Greve & Kappes, 2017)，感知建成环境失序等 (Roccato, Russo & Vieno, 2011)，共同影响安全感。比如，Roccato, Russo & Vieno (2011) 发现，对于有受害经历的居民来说，相比于生活在低建成环境失序社区，生活在高建成环境失序社区，受害经历对不安全感的提升作用会更强。

街景绿视率和居住区统计植被水平既有区别，也有很多联系。据此，本研究提出以下假设：

H2：街景绿视率影响个体安全感知，同时这种影响会被个体直接受害经历调节。

### 5.2.3 街景路灯比例和直接受害经历对安全感知的作用存在交互效应

多个城市街道路灯提升项目的实验研究表明，相对于路灯提升项目开展前，在路灯安装后，行人夜间的不安全感明显下降(Cozens et al., 2003; Cho et al., 2019)。因此，类似假设 H2，本研究提出以下假设：

H3：街景路灯比例与个体直接受害经历存在交互效应，交互项会显著影响个体安全感知。

### 5.2.4 街景绿视率对安全感知的作用受建成环境失序中介，同时这个关系受非正式社会融入调节

前人研究指出，建成环境失序对于安全感知的影响，可能受到社会融入水平

的调节 (Ross & Jang, 2000)。即，在非正式社会融入水平较高的社区，建成环境失序感知对于不安全感的促进作用会被抑制；在非正式社会融入水平较低的社区，建成环境失序感知对不安全感的促进作用则可能会被强化。据此，结合之前假设 1，本研究提出假设 4：

H4：非正式社会融入调节了绿视率通过建成环境失序对不安全感的中介作用。

具体来看，这个假设是指，对于非正式社会融入高的个体，建成环境失序在街景绿视率和不安全感之间的中介效应可能更强。生活在高非正式社会融入和高绿视率水平环境的居民，个体可能感受到的建成环境失序更弱，从而建成环境失序对于不安全感的影响也越弱。相反，对于非正式社会融入低的个体，建成环境失序在街景绿视率和不安全感之间的中介效应可能越弱。

### 5.3 研究方法

#### 5.3.1 中介效应分析

中介论最早在化学领域提出。一些分子中存在电子转移现象，但是不能用传统的结构式来解释，这种转移现象产生的效应就是中介效应。具体是指，根据自变量 X 和因变量 Y 之间的已知关系，探讨 X 是否通过中介变量 M 间接作用 Y。如果该效应显著存在，M 就是中介变量。比如，消费者 (A) 通过外卖平台 (B) 寻找合适的快餐 (C)。那么，外卖平台是中介变量，而外卖平台所起的效果就是中介效应。常用的中介效应检验手段是因果步骤法和系数乘机法。

因果步骤法由 Baron 和 Kenny (1986) 年提出，一般分为三步 (Baron & Kenny, 1986; 温忠麟等, 2004)。

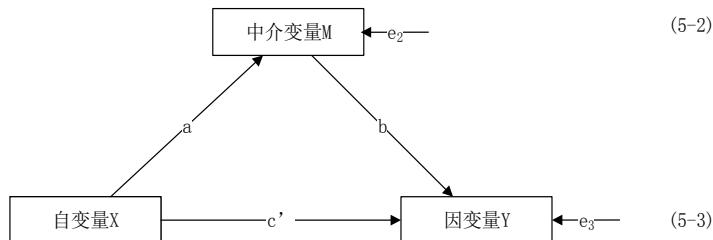


图 5-2 中介效应示意图

$$Y=cX+e_1 \quad (5-1)$$

$$M=aX+e_2 \quad (5-2)$$

$$Y=c'X+bM+e_3 \quad (5-3)$$

第一步，分析等式 (5-1) 的系数值  $c$ ，它是变量  $X$  对  $Y$  的总效应。当  $c$  不显著，拒绝原假设，不进行下一步的中介效应检验。当  $c$  显著，则继续进行下一步的检验。

第二步，分析等式 (5-2) 的系数值  $a$ ，它是变量  $X$  对中介变量  $M$  的作用。当  $a$  显著，表示变量  $X$  会显著作用中介变量  $M$ 。当  $a$  不显著，则需要进一步的 Sobel 检验 (Sobel, 1982)。

第三步，分析等式 (5-3) 的系数  $b$ ，它是中介变量  $M$  对  $Y$  的作用。如果  $a$  和  $b$  都统计意义显著，说明  $M$  的中介效应存在。当  $a$  和  $b$  其中之一不显著，需要进一步的 Sobel 检验。当 Sobel 检验有显著的结果时，中介效应仍然存在。

当达到上述要求，如果等式 (5-3) 中的系数  $c'$  不显著，说明它是完全中介效应。如果系数  $c'$  显著，是部分中介效应。

虽然因果步骤系数法简单有效，但是它的检验效力在所有方法中最低。比如， $A$  对  $B$  有正的  $R$  效应， $B$  对  $C$  有负的  $S$  效应，其实  $B$  在  $A$  和  $C$  之间的关系中起中介作用。但是，由于  $R$  和  $S$  分别是正负项，效果相互抵消，产生遮掩效应。这样，因果步骤法可能计算得出  $A$  和  $C$  之间没有显著的关系，忽视一些现有的中介效应。

在传统因果步骤法基础上，发展出了系数乘积法。系数乘积法舍弃以  $c$  值是否显著为中介效应存在的必要条件，检验中介效应是否显著地不等于 0，利用置信区间的范围来判断中介效应是否存在。系数乘积法的检验效力要优于因果步骤法 (MacKinnon, Fairchild & Fritz, 2007)，目前的大多数研究使用系数乘积法检测

中介效应的显著性。系数乘积法主要包括以下两种：

### (1) Sobel 检验

Sobel 检验法综合考虑系数  $a$  和  $b$ , 即  $a*b$  是否有意义, 从而避免分别检验系数  $a$  和系数  $b$  而导致置信度降低的不足。根据 Sobel (1982), 检验统计值为  $z = ab/s_{ab}$ ,  $s_{ab}$  的计算公式为:

$$s_{ab} = \sqrt{a^2 s_b^2 + b^2 s_a^2} \quad (5-4)$$

$s_b^2$  和  $s_a^2$  分别是  $a$  和  $b$  的标准误。

Sobel 法的检测能力高于因果步骤法 (MacKinnon, Fairchild & Fritz, 2007; 温忠麟等, 2004), 并且可以检测到更多的中介效应。但是, Sobel 法中推导系数乘积的统计量需要假设  $ab$  是正态分布, 这个条件不容易满足。比如, 即使  $a$  和  $b$  是正态分布, 两者的乘积可能不是正态分布。因此 Sobel 检验法也有一定局限性。

### (2) Bootstrap 检验

基于标准误差理论, Bootstrap 方法将大样本作为总体, 执行有放回的抽样以确定更准确的标准误差。Bootstrap 的前提是样本能代表总体, 它的取样方案多种多样。举例, 把一个大小是 1500 的样本估计成 Bootstrap 整个总体, 通过替换进行多次取样, 则能够获得一个大小是 1500 的 Bootstrap 。这些 Bootstrap 样本能够相应计算出 1500 个乘积得分。对得分由低到高排序, 其中 2.5 % 和 97.5 % 构成  $ab$  的置信区间, 置信度为 95%。当置信区间没有零值, 表示系数的乘积显著, 有间接效应 (Preacher & Hayes, 2008; Preacher, Rucker, & Hayes, 2007; 温忠麟, 刘红云& 侯杰泰, 2012; 温忠麟, 叶宝娟, 2014)。当置信区间存在零值, 没有显著的间接效应。Bootstrap 法具有比 Sobel 法等其他中介效应检验方法更准确的检测力, 目前广泛应用于中介效应分析。

## 5.3.2 调节效应和交互效应分析

调节效应是检验第三个变量是否影响自变量和因变量管理的强度和方向的方法 (James & Brett, 1984), 概念模型见图 5-3。调节变量可以是二值化变量如性别等, 连续变量如降雨量或定距变量年龄等。

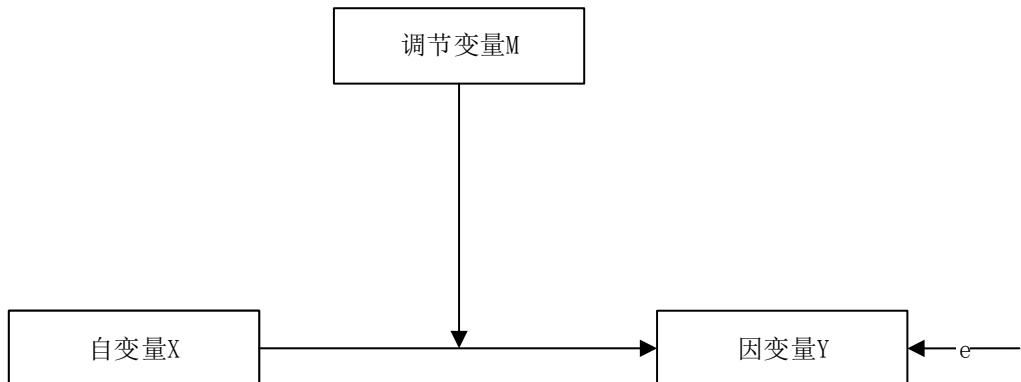


图 5-3 调节变量示意图

根据调节变量和自变量的类型，一般选取不同的调节检验方法（侯杰泰，温忠麟 & 成子娟，2006）。当自变量和调节变量属于连续型变量，则选择带乘积项的回归方程，使用层次回归分析法完成调节效应检验。方法是，先对自变量 X 和调节变量 M 做 Y 的回归，得到系数值  $R_1^2$ 。然后，分析 Y 对 X、M、和乘积 XM 的回归，得到系数值  $R_2^2$ 。当  $R_2^2$  统计意义上地超过  $R_1^2$ ，表明调节效应存在。如果情景中是调节变量由分类变量构成，自变量由连续型变量构成，那么最好使用分组回归方法。其他的选取规则见下表。

表 5-1 调节效应选取规则

调节变量	自变量	
	分类变量	连续变量
分类变量	ANOVA（方差分析）法	分组回归法或分层回归法
连续变量	分层回归分析法	分层回归分析法

交互效应与调节效应类似。A 影响 C，B 影响 C，则 AB 的乘积既可以计算出 B 在 A 对 C 的影响上的调节作用，也可以计算出 A 和 B 共同作用于 C 的交互作用。即使 A 不影响 C，B 影响 C，如果 AB 的乘积项显著，可以表明，A 虽然不直接影响 C，但是通过和 B 的交互作用而影响 C。

### 5.3.3 有调节的中介效应分析

在一些情景下，调节效应和中介作用会一起出现，即可能出现有中介的调节效应和有调节的中介效应。本研究关注有调节的中介效应。有调节的中介模型重点是中介模型，然后是这个中介模型被调节，在不同情景下（调节变量的不同水平下）有不同的中介效果。有调节的中介效应由 Muller, Judd & Yzerbyt (2005) 和 Preacher, Rucker & Hayes (2007) 分别提出。

根据不同情景，有调节的中介效应模型有如下几种类型（图 5-4）。

模型 A，该模型中，一个自变量同时调节中介变量和因变量之间的连接。模型 B，该模型中，调节变量调节自变量和中介变量之间的连接。模型 C，该模型中，一个调节变量调节中介变量和因变量间的连接。模型 D，该模型中，同一个调节变量既调节自变量和中介变量间的连接，也调节中介变量和因变量之间的连接。模型 E，该模型中，一个调节变量影响自变量和中介变量之间的连接，另一个调节变量影响中介变量和因变量之间的连接。

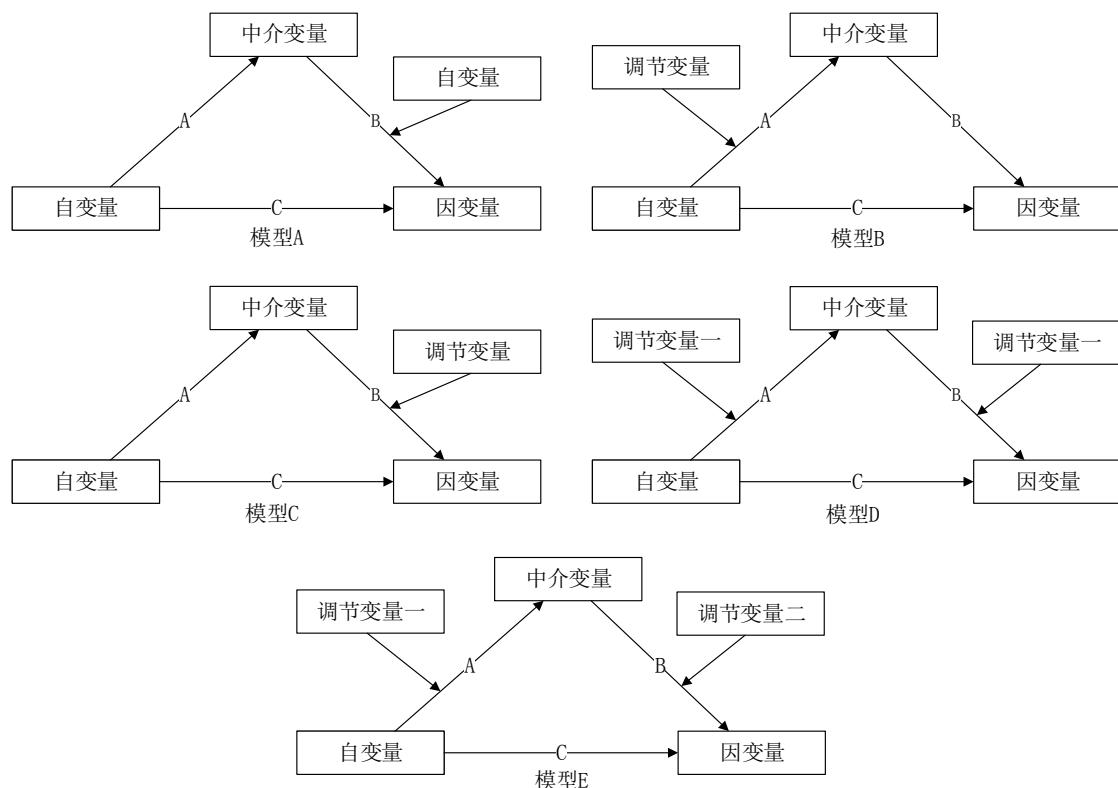


图 5-4 有调节的中介效应模型示例

#### 5.3.4 变量设定

##### (1) 因变量

本文的因变量为基于问卷数据的居民安全感知指标，以“不安全感”为因变量。不安全感指标与第五章一致，由被抢风险感知、被偷风险感知等五项指标的均值组成。

##### (2) 自变量

不同模型中选取的自变量不同。自变量包括街景绿视率和街景路灯比例。

### (3) 中介变量

本文选择基于问卷数据的感知社区环境作为中间变量，具体为感知的建成环境失序指标。

### (4) 调节变量

本文选择之前基于问卷数据的直接受害经历指标作为研究的调节变量。

### (5) 协变量

将第五章中与因变量显著相关的其他变量加入作为协变量，包括性别、间接受害经历、感知的社会环境失序、感知的社区警力、社区教育水平、社区商业设施比、社区政府机构比等指标。

## 5.4 研究结果

### 5.4.1 中介效应结果

本研究将 Hayes (2018) 开发的 PROCESS v.3.5 插件置入 IBM SPSS Statistics 22，选取 MODEL 4 为基本的中介效应模型，置信区间设置为 95%，Bootstrap 样本量为 5000 次，使用 Bootstrap 方法计算置信区间，所有变量都中心化。完成相关设置。

由表 5-2 可知，街景绿视率对不安全感的总效应是 -1.132，对应的 95% 置信区间 [-1.771 - -0.493]，不包含 0，表明总效应存在。街景绿视率对不安全感的直接效应是 -0.857，对应的 95% 置信区间为 [-1.485 - -0.229]，不包含 0，表明直接效应存在。街景绿视率经建成环境失序对不安全感的间接效应为 -0.275，对应的 95% 置信区间为 [-0.442 - -0.132]，不包含 0，表明间接效应存在。根据以上分析结果，建成环境失序在街景绿视率与居民不安全感之间发挥部分中介效应。假设 H1 得到验证。图 5-5 是该中介效应分析的图形结果。

表 5-2 建成环境失序在街景绿视率-不安全感关系中的中介效应分析

	Total Effect	SE	t	p	LLCI	ULCI
街景绿视率-不安全感	-1.132	0.326	-3.476	0.001	-1.771	-0.493
街景绿视率-不安全感	Direct Effect	SE	t	P	LLCI	ULCI
-0.857	0.320	-2.676	0.008	-1.485	-0.229	
Indirect Effect	BootSE	BootLLCI	BootULCI			
街景绿视率-建成环境失序 -不安全感	-0.275	0.079	-0.442	-0.132		

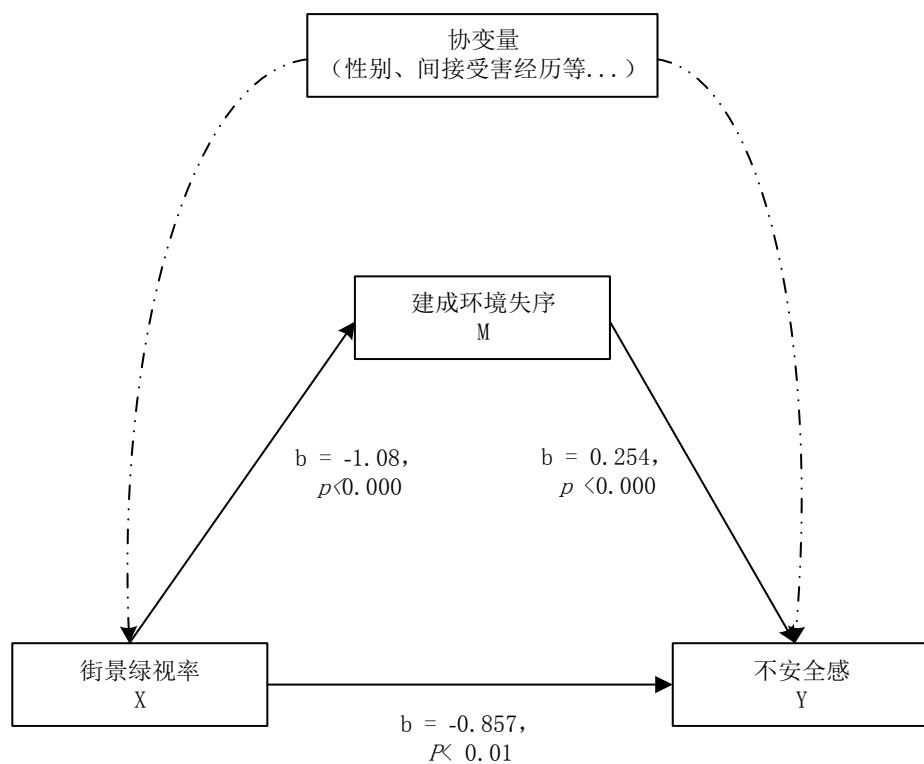


图 5-5 中介效应可视化结果

#### 5.4.2 调节效应和交互效应结果

在调节效应分析前面, 需要对相关的自变量和调节变量做中心化处理, 这样得到的交互项不会和自变量、调节变量产生高度的相关。本研究也选取 Hayes (2018)开发的 PROCESS v.3.5 插件置入 IBM SPSS Statistics 22, 选取 MODEL 1 为基本的调节效应模型, 把所有因子进行中心化处理, 完成基本设置。

关于直接受害经历是否和街景绿视率、街景路灯比例之间存在调节或交互效

应如表 5-3 所示。从表中可以看出，在未加入调节项时（Model 1），街景绿视率和直接受害经历都是显著地影响居民不安全感（分别为 Coeff=-0.098, P <0.01; Coeff=-0.408, P<0.01）。在考虑调节效应的模型中，Model 2 的结果显示，直接受害经历和街景绿化的交互项是统计意义较为显著的变量（Coeff=-1.02, P <0.1），方向为负，说明直接受害经历强化了街景绿视率对于不安全感的减弱作用。研究假设 H2 得到证明。

此外，在未加入交互项时（Model 1），街景路灯比例没有显著地影响居民安全感（p>0.1）。在考虑交互效应的模型中（Model 3），直接受害经历显著影响不安全感，街景路灯比例和直接受害经历的交互项也在 0.01 水平上显著地影响居民不安全感（Coeff=-1.217, P <0.01），方向为负，说明街景路灯比例通过直接受害经历影响不安全感。研究假设 H3 得到证明。

表 5-3 调节效应和交互效应检验结果

不安全感						
	Model 1		Model 2		Model 3	
	Coeff.	SE	Coeff.	SE	Coeff.	SE
截距	2.060**	0.144	2.557**	0.023	2.088**	0.116
街景绿视率（X <sub>1</sub> ）	-0.098**	0.348	-1.001**	0.320		
街景路灯比例（X <sub>2</sub> ）	-0.089	0.178			0.133	0.177
直接受害经历（W）	0.408**	0.048	0.404**	0.048	0.16**	0.057
街景绿视率×直接受害经历			-1.02#	0.659		
街景路灯比例×直接受害经历					-1.217**	0.363
R <sup>2</sup>	0.119		0.121		0.124	
F	26.086**		23.772**		23.466**	

注：\*, p<0.05; \*\*p<0.01; #, p<0.1. 省略协变量因素的结果。

根据研究结果，本研究进一步得出调节效应图。从图5-6可知，两条曲线的斜率不同，验证了受害经历在街景绿视率和不安全感关系中的调节效应，以及在街景路灯比例和不安全感关系中的交互效应。当有受害经历时，线的歪斜水平变大，不安全感的下降程度加快；当无受害经历时，线的歪斜水平变缓，不安全感的下降程度变缓。

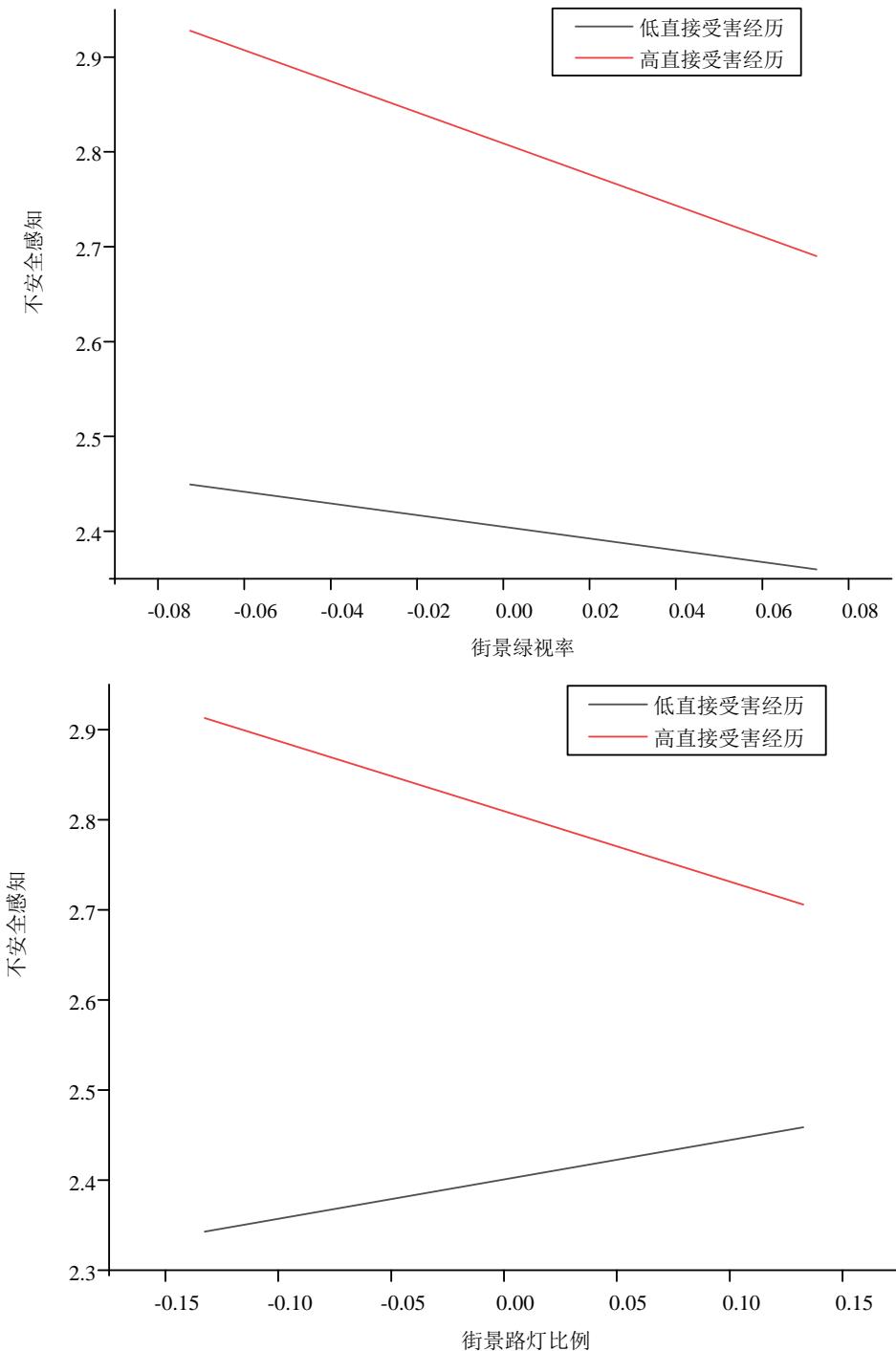


图 5-6 调节效应图

(注：横坐标街景绿视率值存在正负的区别，因为变量在进入模型前进行了中心化操作；直接受害经历也进行了中心化操作，分别以25%和75%位的值为低值和高值)

### 5.4.3 有调节的中介效应结果

进行有调节的中介效应分析，基本的参数设置和中介效应分析一致，具体选择模型有所差别。本研究选取 Hayes (2018) 开发的 PROCESS v.3.5 插件中的 MODEL 58 为标准的有调节的中介效应模型，把所有变量做中心化处理，完成相

关设置。如表 5-4 所示, 绿视率显著影响了居民的建成环境失序感知 ( $\text{Coeff}=-1.063, P<0.00$ ), 而且非正式社会融入显著强化了绿视率对居民建成环境失序感知的负向作用 ( $\text{Coeff}=-0.658, P<0.05$ )。也即, 社区绿视率的提高会降低居民的建成环境失序感知, 同时社区非正式社会融入水平的增加会强化这种趋势。此外, 绿视率显著影响了居民的不安全感 ( $\text{Coeff}=-0.849, P<0.01$ ), 建成环境失序也显著影响了居民的不安全感 ( $\text{Coeff}=0.254, P<0.01$ ), 非正式社会融入显著强化了建成环境失序对居民不安全感的正向作用 ( $\text{Coeff}=0.095, P<0.01$ )。

因为街景绿视率的直接效应同时存在, 因此为一个部分中介模型。进一步, 当非正式社会融入较低时, 建成环境失序的中介效应为-0.097, 但是置信区间跨越 0 值, 不显著; 当非正式社会融入中等时, 建成环境失序的中介效应为-0.27, 同时置信区间没有跨越 0 值, 表明中介效应显著; 当非正式社会融入高时, 建成环境失序的中介作用为-0.521, 置信区间没有跨越 0 值, 中介效应显著。因此, 建成环境失序的中介效应随着非正式社会融入的增加而增强。综上。经过实验, 发现街景绿视率对安全感知的作用受建成环境失序中介, 同时受非正式社会融入调节, 即研究假设 H4 得到证明。有调节的中介效应的可视化见图 5-7。

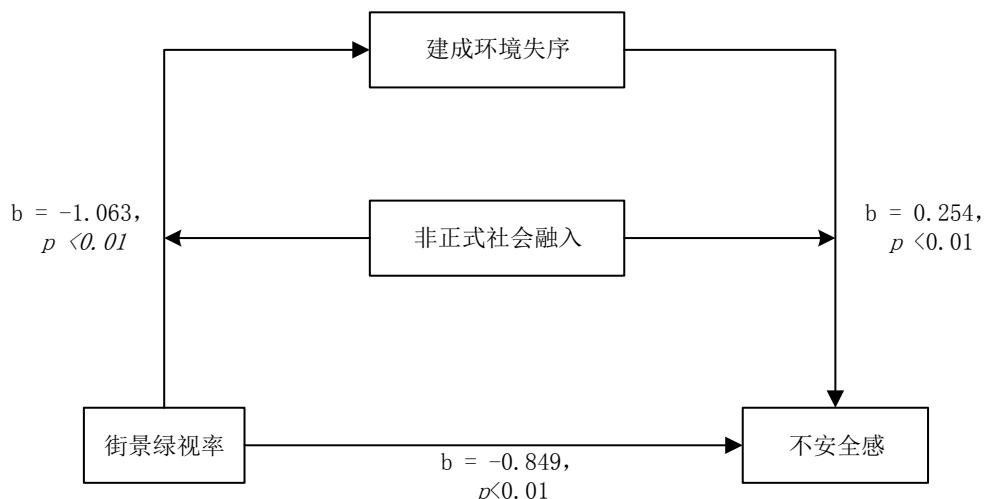


图 5-7 有调节的中介效应图

表 5-4 社会融入调节街景绿视率-建成环境失序-不安全感中介效应模型回归结果

以建成环境失序为因变量						
	系数	标准误	T值	P值	下限 (95%)	上限 (95%)
常数	-0.001	0.019	-0.071	0.943	-0.039	0.036
街景绿视率	-1.063	0.264	-4.02	0.000	-1.581	-0.544
非正式社会融入	0.097	0.024	4.002	0.000	0.049	0.144
街景绿视率× 非正式社会融入	-0.658	0.318	-2.071	0.039	-1.281	-0.035

$R^2=0.021$ ;  $R^2\text{变化}=0.002$ ;  $F=4.288^*$

以不安全感为因变量						
	系数	标准误	T值	P值	下限 (95%)	上限 (95%)
常数	2.553	0.023	110.576	0.000	2.507	2.598
街景绿视率	-0.849	0.319	-2.664	0.008	-1.475	-0.224
建成环境失序	0.254	0.029	8.795	0.000	0.197	0.311
非正式社会融入	-0.084	0.029	-2.857	0.004	-0.141	-0.026
建成环境失序× 非正式社会融入	0.095	0.029	3.26	0.001	0.038	0.151

$R^2=0.06$ ;  $R^2\text{变化}=0.006$ ;  $F=10.628^{**}$

不同非正式社会融入条件下的建成环境失序的中介效应和有调节的中介效应的判定指标				
调节变量	建成环境失序的 中介效应	BootSE	Boot下限	Boot上限
			(95%)	(95%)
非正式社会融入低 (-1SD)	-0.097	0.082	-0.28	0.042
非正式社会融入中等	-0.27	0.078	-0.433	-0.129
非正式社会融入高 (+1SD)	-0.521	0.153	-0.836	-0.232

注: \*\*  $p < 0.01$ ; \*  $p < 0.05$ .

## 5.5 小结

本章节关注街景微环境对居民不安全感的间接效应。以社区失序理论为理论基础, 结合已有研究成果, 本章节提出了街景绿视率对不安全感的主效应、建成环境失序的中介效应、受害经历的调节效应、街景路灯比例的交互效应等假设。通过对收集数据的分析和检验, 明确了各变量之间的关系, 证明了在非正式社会融入的调节下, 街景绿视率经由建成环境失序感知引发居民不安全感的作用路径, 丰富了客观社区环境对安全感知影响的作用机理研究。本研究的结论可以为下一

章节的个体安全感知预测的输入变量选择和建模方法选择提供理论支撑。主要的研究结论如下：

(1) 安全感知是由包含调节效应、交互效应和中介效应在内的复杂机制驱动

本研究验证了客观微环境因素，感知环境因素，个体因素存在相互联系的通道，以中介作用或调节作用的方式最终作用于安全感知。因此，安全感知的形成机制较为复杂。

(2) 微环境要素对安全感知不仅存在直接效应，还存在间接效应

本研究验证了客观绿视率微环境不仅对安全感知有直接作用，而且和直接受害经历以及非正式社会融入程度产生调节作用。因此，客观微环境因素可能不仅直接影响安全感知，而且可能通过其他个体因素或环境因素影响安全感知。

(3) 客观环境对于主观安全感知的影响机制受感知的环境因素中介

与之前的文献一致，本研究验证了客观环境对于主观安全感知的影响是通过感知的环境因素来作用这一假设。不同于之前文献，本研究使用了基于街景图片的一种新型环境指标，即街景绿视率。由于居民的日常活动是基于道路开展的，这两个指标更加真实地刻画了居民的日常活动环境暴露。中介效应分析发现，街景绿视率影响了居民的建成环境失序感知，居民的建成环境失序感知影响了居民的不安全感。同时，街景绿视率影响了居民的不安全感。Bootstrap 检验表明，间接效应显著，且是部分中介效应。这个发现丰富了环境因素和安全感知关系的理论研究，从理论上支撑了客观环境因素能够作用居民的安全感知。

(4) 直接受害经历和客观环境在主观安全感知的影响机制中起到调节作用和交互作用

本研究发现，直接受害经历会强化绿视率对安全感知的抑制作用。即，对于个体而言，有受害经历，会导致街景绿视率对于不安全感的抑制作用增强。这与 Rühs, Greve & Kappes (2017) 的研究结论一致。他们的研究认为，当个体遭遇受害经历等创伤事件后，会主动去寻求找资源进行恢复，同时也会进行自我调节，从而缓解或消除被害经历对个体带来的负面影响。因此，有受害经历的人，可能更加关注周围环境的线索 (Jing et al., 2020)，以及更加会关注周围的自然环境，访问街道绿色空间的次数可能更多，从而街景绿视率水平对这类群体的不安全感作用比无受害经历的群体更加明显。还有一种可能原因，有受害经历的个体，更可能寻求帮助及恢复，如通过访问公园等开放空间缓和心情。Van den Berg et

al.(2010)的研究,证明了在遭受创伤事件后,居住区绿地可以有利于个体的健康水平提升。因此,绿视率高,对于有受害经历的个体可能更加有用。同时,虽然街景路灯比例没有显著影响居民的安全感知,但是生活在街景路灯比例高的社区,有直接受害经历的人的不安全感下降程度更快,也说明了街景路灯比例对安全感知存在间接作用。在受到受害经历等创伤事件后,社区街景路灯比例作为环境恢复的一个载体,有利于抑制受害经历对居民安全感知的负面作用。

(5)“客观环境-感知环境-安全感知”的作用机制中既存在中介作用,也存在调节作用

本研究分析表明,在“客观环境-主观环境-安全感知”的作用机制过程中,既存在中介效应,也存在调节作用。特别地,非正式社会融入在“街景绿视率-建成环境失序-安全感知”中介效应机制中产生显著的调节效应。其中,社区绿视率的提高会降低居民的建成环境失序感知,同时社区非正式社会融入水平的增加会强化这种趋势,这与预期一致。社区建成环境失序水平的增加会强化居民的不安全感,而社区非正式社会融入水平的增加会强化这种增加不安全感的趋势。这是由于,社区非正式社会融入水平虽然可能抑制不安全感,但是可能也会导致居民知晓更多间接受害经历,从而促进不安全感的上升(Gates & Rohe, 1987; Kanan & Pruitt, 2002)。可知,感知环境因素和建成环境因素之间存在交互作用。安全感知感受到两类环境因素的共同影响。

环境对于安全感知的影响存在一个复杂的机制。本章节的机制分析探索了“客观环境-主观环境-安全感知”的影响过程。虽然只是理论上的分析,但是有一定实际作用。首先,社区居民的安全感知是主客观环境共同作用的结果,客观环境通过主观环境感知起重要作用。在提升社区居民安全感的实际工作中,各个领域的专家应该协调。既要运用环境设计方法来进行预防犯罪和抑制不安全感,也需要通过消除贫困等方法来从社会根源上消除居民的社会不安全状态。同时,针对不同类型的人群,针对不同类型的社区,需要因地制宜,采取不同的策略,比如根据不同类型的社区开展不同的环境设计项目和社区集体活动,来提升社区居民的安全感。比如,在城市规划和建设中,要增强城市的美丽度和绿化,促进人们,特别是弱势群体的安全感提升。



## 第六章 居民安全感知预测及社区安全感知模拟

### 6.1 引言

本研究围绕居民安全感知展开。前述章节重点开展了环境要素刻画、影响因素分析、机理分析等工作。在前述章节的基础上，本章节着重开展安全感知的预测及空间模拟，以期为实际的政策干预及事后评估提供方法指导。

本章节的主要研究目的有两方面。其一，研究是否非线性模型比线性模型对个体安全感知有更好的预测能力。在前述章节中的影响因素分析中，研究关注点在于相关因素与安全感知的关系是否显著，线性回归分析方法可以直接得出各影响因素与安全感知的关系程度，因此线性回归方式是研究影响因素的合理选择方法。但是，线性回归方法对于安全感知模型的预测能力是否最优，还有待进一步研究。安全感知作为一种主观情感感知，可能存在不同的阈值区间触发不同程度的安全感知，这种变化趋势可能是非线性的。而一系列非线性回归方法，如支持向量机、决策树、人工神经网络等算法，目前已经在社会科学中广泛运用，用以分析各种社会问题。因此，本章节尝试分别建立非线性和线性的安全感知预测模型，比较非线性模型和线性模型的模型效果差异。

其二，本章节尝试开发一套安全感知的空间模拟方法。在实际城市治理中，政府部门为了获取某地区居民的安全感知状况，通常是随机发放批量问卷，然后通过问卷回收，计算得出该地区居民的平均安全感知值。但是，这种方法存在成本高、难以在大范围开展、时效性低等缺点。当今社会的信息化进程加快，各类城市数据资源十分丰富。本章节尝试基于个体问卷样本，耦合社区尺度的多源数据，得到个体安全感知的预测模型，然后对广州市所有居民个体进行安全感知预测。并将结果聚合到社区尺度，得到社区尺度的安全感知，并完成空间制图。

本章节的研究具有理论和实际意义。理论上，之前还没有研究比较非线性和线性方法对于安全感知模型的预测效果差异，也没有研究探讨实现一种城市区域范围内高时空精度安全感知的空间模拟方法。因此，在安全感知模拟领域，本章节的研究具有方法上的创新意义。实际上，本研究对于城市规划和管理具有指导作用。通过高精度的安全感知时空分布制图，可以有效指导政府部门的资源调配。

图 6-1 是社区安全感知空间模拟的流程图。本研究的主要步骤是先对比分析得出最优的个体安全感知预测模型，然后利用预测模型预测整个研究区范围内个体的安全感知水平，最后聚合到社区尺度上，进行可视化分析。主要步骤如下：

(1) 确定安全感知。本研究的结果变量为安全感知，原始问卷调查中的安全感知是连续变量。通常来看，二值化的结果变量可以使预测任务更加简单，并取得较好的预测效果。考虑到安全感知作为一种情感感知本身具有难以预测的特点，而本研究的任务是进行社区尺度的安全感知空间模拟。因此，本研究将安全感知进行二值化处理（0 为有安全感，1 为无安全感）。

(2) 个体的安全感知影响因素构建。由于本研究的最后步骤是利用预训练模型预测整个研究区个体的安全感知，而整个研究区个体的个体属性拟基于统计数据模拟生成。因此，个体安全感知预测模型构建时需要考虑运用和统计数据一致的个体变量。经过理论分析和实际可操作性分析，本研究选取个体的年龄、性别、受教育程度、户口等四个个体变量作为个体属性变量。个体所属社区的社区指标与前述章节的社区环境指标一致，包括街景指标（路灯比例、绿视率、天空开阔指数、垃圾度）、POI 指标（商业点比例、政府机构点比例）、遥感指标（NDVI 指数）、犯罪率和社区统计指标（受教育水平、外来人口比例和人口密度）等。综合来看，一共 15 个指标为特征变量，作为预测安全感知模型的输入变量。

(3) 个体的安全感知影响因素确定。在这些输入变量中，个体属性指标是固有指标，时间分辨率较低。社区环境指标中，街景指标、POI 指标、遥感指标、犯罪率指标等时间分辨率较高，而社区统计指标来源于六普统计数据，十年更新一次，时间分辨率较低。为了实现时效性较高的安全感知预测模型，本研究拟将社区统计数据指标的时间分辨率提高。利用中国联通在研究区内 2016 年特定时段的手机数据，耦合社区统计数据，得到基于手机数据的社区统计指标，可以更精准地刻画问卷同时期的社区统计指标。因此，通过耦合手机数据和统计年鉴数据，本研究将模型的输入变量进行更新，并确定最终的安全感知预测模型的输入变量。

(4) 将基于多源数据集获取的输入变量分别放置进入 1 个广义线性回归模型（二元逻辑回归模型）和 2 个非线性模型（支持向量机模型和随机森林模型），通过对不同模型进行结果精度比较，选取最优模型。

(5) 根据研究区内的广州市 2010 年统计年鉴的 1753 个社区基本结构数据

(包括性别人数、各类年龄人数、各类受教育人数、各类户口人数 4 个社区指标), 模拟出符合社区统计数据比例的 9328963 条个体数据。然后, 将每个个体所属社区的其他 11 个社区环境数据赋予到个体上。最终, 每条个体数据包括 4 个人口统计特征数据(个体性别、个体年龄、个体受教育程度、个体户口状态)和 11 个社区各项环境指标数据。

(6) 运用训练效果最优的模型对 9328963 条个体数据进行个体安全感知的预测。按照所属社区原则, 将个体安全感知结果以均值的方法聚合到社区, 完成社区安全感知的计算, 对运算结果进行精度验证。最后, 绘制出社区安全感知的空间分布制图结果。

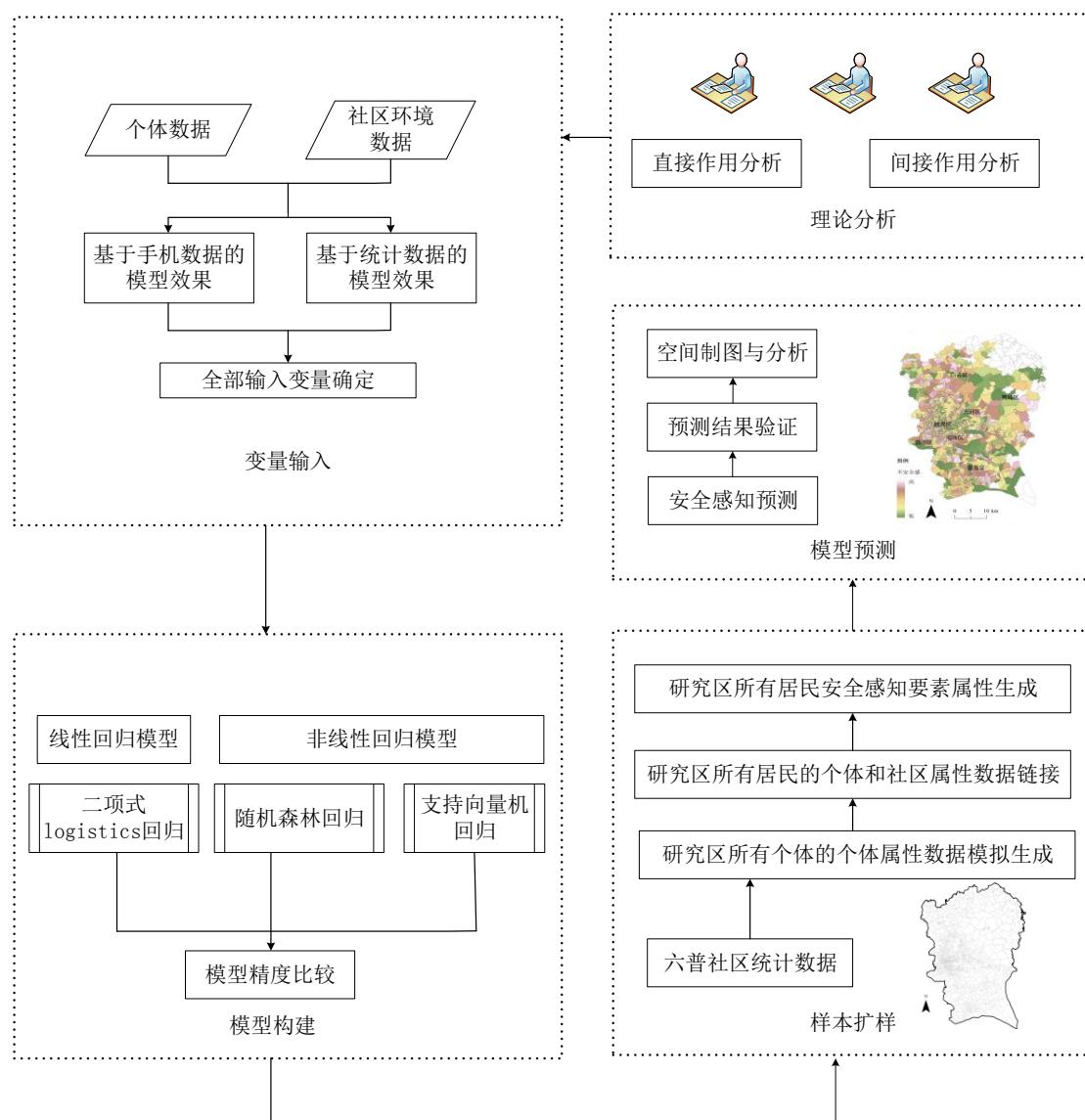


图 6-1 社区安全感知模拟流程图

## 6.2 研究方法

### 6.2.1 线性和非线性回归模型设定

#### 6.2.1.1 二项式 logistics 回归模型

根据预测变量 X 预测响应变量 Y 的时候，如果 Y 是连续型变量，一般使用多元线性回归分析。如果 Y 是定类数据，一般使用 logistics 回归分析。Y 作为定类数据可以分为二元定类变量，多元有序变量，多元无序变量，对应具有二项式 logistics 回归，多元有序 logistics 回归，多元无序 logistics 回归。本章节的因变量为二分变量，选用二项式 logistics 回归。

传统线性模型如公式 4-1（第四章）所示。对于二项式 logistics 回归模型，Y 的取值为 0 或 1，取值区间为 (0,1)。而回归方程 4-1 右边部分的结果值是连续的，取值区间在  $(-\infty, +\infty)$ 。这种情况下，如果要让方程 4-1 等式成立，需要对公式左边或右边做一个变化，使二者的取值区间相同。一个简单的方法就是对左边 Y 值进行一次数学 logit 变化，这样 Y 的取值区间就为  $(-\infty, +\infty)$ 。具体做法就是通过二项分布的方法，把 0, 1 两个值与概率 P 和 1-P 对应起来（记值 1 出现的条件概率为 P, 值 1 未出现的概率为 1-P），构造一个事件发生比  $Odds = \frac{P}{1-P}$ ，Odds 的变化范围是  $(0, +\infty)$ 。因为公式右边回归方程范围是  $(-\infty, +\infty)$ ，还需要进行一次变化，这个变化通过对数函数 log 实现。

这样，可以得到式 6-1：

$$\log\left(\frac{P}{1-P}\right) \in (-\infty, +\infty) \quad (6-1)$$

进一步，建立如下回归方程 6-2，

$$\log\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \mu \quad (6-2)$$

公式 6-3 就是 logistics 回归方程，每个样本与一个概率值 P 对应，通过简单变化，可得到 P 的计算公式，

$$P_n = \frac{e^{\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \mu}}{e^{\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \mu} + 1} \quad (6-3)$$

进一步，根据最大似然估计（Maximum Likelihood Estimation）的方法可以计算出各参数值  $\beta_1, \beta_2, \dots, \beta_n$ 。

### 6.2.1.2 随机森林模型

随机森林是一种用途在分类、回归或其他任务的集成学习法，是在训练时构造很多个决策树并输出类，该类是单个数的类（分类）或均值（回归）的模式。随机决策森林修正了决策树经常过度拟合训练集的不足。Ho (1995) 使用 Random Subspace Method 首先提出了随机决策森林算法。该算法由 Breiman (2001) 和 Cutler & Zhao (2001) 扩展。这个扩展结合了 Breiman 的“bagging”思想和由 Tim Kam Ho 提出的特征随机选择理念。

#### (1) 决策树

因为随机森林模型是由很多棵决策树组成的，有必要首先简要描述决策树。举例，面临一个问题：“是否应该去目的地 A 吗？”从这个问题开始，可以问一系列问题来确定答案，比如，“这个路程时间长吗？”“去目的地 A 的成本？”这些问题构成了树中的决策节点，作为分割数据的手段。每一个问题都帮助个体做出一个最终的决定，这个决定用叶节点表示。符合标准的观测会进入 Yes 分支，不符合标准的观测会进入备用路径。决策树寻求发现数据子集的最佳分割，通常通过分类和回归树（CART）算法进行训练。一些指标，诸如 Gini Impurity 基尼系数杂质度，Information Gain 信息增益，或 Mean Square Error (MSE) 均方误差，可以用来评估分割的质量。

#### (2) 集成方法

集成学习方法由一组分类器（如决策树）组成，它们的决策被聚合起来以识别出最有可能的结果。Breiman(1996)介绍了“bagging”方法。这个方法就是指，在一个训练集中选择一个随机的数据样本作为替换，这意味着可以多次选择单个数据点。在生成了几个数据样本之后，然后独立地训练这些模型，根据任务的类型，即分类或回归，这些预测的平均值或大多数会产生更准确的估计。这个方法通常被用来在有噪声的数据集中减少方差。

#### (3) 随机森林算法

随机森林算法是 bagging 方法的扩展，它同时利用 bagging 和特征随机性来创建一个不相关的决策树森林。特征随机，也被称为特征袋或随机子空间方法，生成特征的随机子集，这确保了决策树之间的低相关性。这是决策树和随机森林之间的关键区别。决策树考虑所有可能的特征分割，而随机森林只选择这些特征

的一个子集。通过考虑数据中的所有签证变异性，我们可以减少过拟合、偏倚和总体方差的风险，从而获得更精准的预测。

随机森林算法由一组决策树组成（图 6-2），而集成中的每棵树都由一个从带有替换的训练集中抽取的数据样本组成，称之为 Bootstrap 样本。在这个训练集中，三分之一作为测试数据保留下来，被称为 out-of-bag(oob)袋外数据样本。然后通过特征套袋注入另一个随机性实例，为数据集增加更多的多样性，减少决策树之间的相关性。根据问题的类型，预测的测度将有所不同。对于一个回归任务，个体决策树将被平均，而对于一个分类任务，多数投票即最频繁的分类变量将产生预测类。最后，oob 样本用于交叉验证，进行确定预测。实际实验中，随机森林算法有三个主要的超参数，需要在训练前设置。这些参数包括节点大小、树的数量和抽样特征的数量。通过这个阶段的设置，随机森林分类器可以被用来解决回归或分类问题。

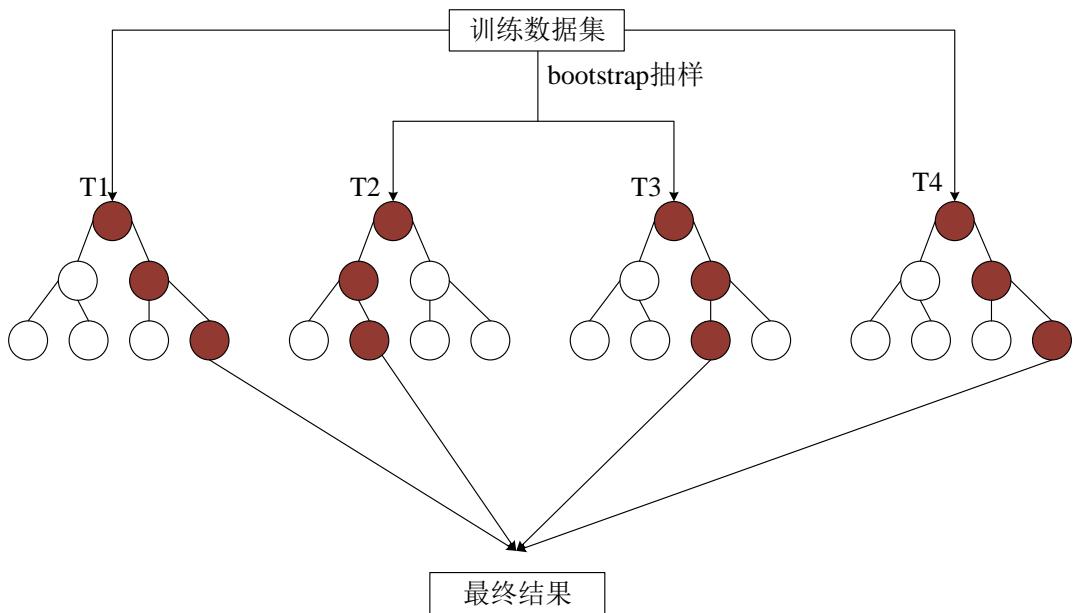


图 6-2 随机森林分类器图

当面临回归或分类问题时，随机森林算法有许多优点和一些不足。优点：

- (1) 减少过拟合的风险。决策树存在过拟合风险，因为它们倾向于紧密拟合训练数据中的所有样本。然而，当随机森林中决策树的数量具有鲁棒性时，分类器不会对模型进行过拟合，因为不相关树的平均降低了总体方差和预测误差。
- (2) 提供灵活性。由于随机森林能够以较高的精度处理回归和分类任务，它在机器学习领域被普遍使用。特征套袋还使随机森林分类器成为估算缺失值的有效工具，因为它在部分数据缺失时保持了准确性。

(3) 容易确定特征的重要性。随机森林使得很容易评估模型中变量的重要性或贡献。有几种方法可以评估功能的重要性。基尼系数重要性(Gini importance)和平均杂质减少(MDI)通常用来衡量排除给定变量后模型的精度降低了多少。然而，排列重要性，也被称为平均下降精度(MDA)，是另一个重要的度量。MDA 通过随机排列 oob 样本中的特征值来识别准确率的平均下降。计算公式为：

$$M(p^j) = \frac{1}{N} \sum_{i=1}^t (e_i^j - e_i) \quad (6-4)$$

其中，变量  $M(p^j)$  的加入与否影响的 oob 误差越大、准确度下降的越快，表示  $M(p^j)$  对于模型越很重要。

不足主要表现在以下三点。第一，耗时过程不小。随机森林算法可以处理大型数据集并提供更加准确的预测，但是在处理数据时可能会比较慢，因为它是为每棵决策树计算数据。第二，需要更多的资源。由于随机森林处理更大的数据集，它们将需要更多的资源来存储数据。第三，更复杂。与众多决策树相比，单个决策树的预测会更容易解释。

### 6.2.1.3 支持向量机（SVM）模型

支持向量机(Support Vector Machine, SVM)为一类有监督的机器学习算法，可以用于分类或回归任务，不过最多是使用于分类任务中。Boser, Guyon & Vapnik (1992)提出了一种通过将核技巧应用于最大间隔超平面来创建非线性分类器的方法。而当前标准的前身(软边距)由 Cortes & Vapnik (1995) 提出。

SVM 算法里，每个数据项被描绘为 n 维空间中的一个点。n 代表拥有的特征数量)，每个特征的值为表示为一定坐标值。然后，通过找到能够很好地区分两个类的超平面来进行分类。其中，支持向量仅仅是单个观测的坐标。支持向量分类器是最能区分两类(超平面/直线)的前沿分类器。

关于如何识别正确的超平面，以区分两个类，可以从多个场景理解。对于明显不同的两个类别，超平面很容易划分。但是，一些场景中，两个类别之间不能划分出线性超平面(图 6-3A)。SVM 通过引入额外的特性解决这个问题，即增加一个新特征  $z=x^2+y^2$ 。从而可以画出 x 轴和 z 轴上的数据点，如图 6-3B 所示。图中，z 的所有值始终为正，因为 z 是 x 和 y 的平方和。在原始图中，圆圈出现在靠近 x 和 y 轴的原点的位置，从而导致 z 的值较低，而星星相对于远离原点，导

致 z 值较高。

另外，SVM 算法具有一种内核技巧（the kernel trick）的技术可以实现自动划分超平面的工作。SVM kernel 是一个占用低维输入控件并将其转换为高维空间的函数，即它将不可分离的问题转换为可分离的问题。它在非线性分离问题中最有用。如图 6-3C，是使用核函数划分出的一个超平面。

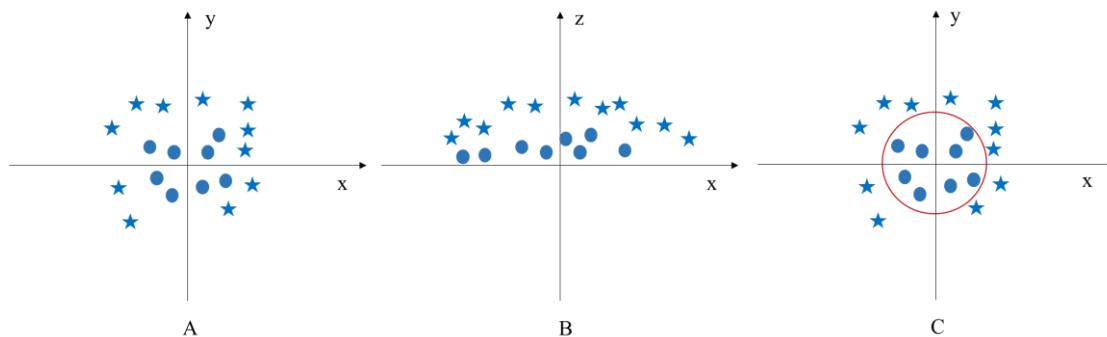


图 6-3 SVM 超平面划分

确定核函数的一般方法是实验多种核函数和参数，最终确定一个分类误差最低的核函数和参数。核函数中，两个关键参数是 GAMMA( $\gamma$ ) 和惩罚因子 (C)，这两个参数的值会直接决定 SVM 模型的分类精度。

核函数主要包如下几种，分别是：

线性核函数

$$K(x_i, y_j) = (x_i, y_j) \quad (6-5)$$

P 阶多项式核函数

$$K(x_i, y_j) = ((x_i, y_j) + 1)^p \quad (6-6)$$

径向基核函数（高斯核）

$$K(x_i, y_j) = \exp\left\{-\frac{\|x_i - y_j\|}{2\sigma^2}\right\} \quad (6-7)$$

SIGMOID 核函数

$$K(x_i, y_j) = \tanh(b(x_i, y_j) + C) \quad (6-8)$$

本文以非线性径向基函数为核，即以高斯核(Radial basis function, RBF)建立支持向量机模型。使用内核函数不需要每次都全部计算出原始样本点映射的新的无限维采样点，而可以使用映射后的新的采样点的点乘计算公式，这减少计算量，也节省存储空间。

## 6.2.2 评价指标

关于个体安全感知预测的不同模型效果对比，本章节主要通过混淆矩阵的相

关指标来检验。混淆矩阵的构建原理如下表 6-1。基于混淆矩阵的常用评价指标见表 6-2。

表 6-1 混淆矩阵原理

混淆矩阵		真实值	
预测值		Negative	Positive
		Negative	TN
		Positive	FP(Type II)      TP

表 6-2 基于混淆矩阵的常用评价指标

评价指标	公式	意义
准确率	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$	全部预测正确结果在全部预测量中所占的比例
精确率	$Precision = \frac{TP}{TP + FP}$	判断为 Positive 的全部值中，算法判断正确的比例
召回率	$Recall = \frac{TP}{TP + FN}$	实际值为 Positive 的全部值中，算法判断正确的比例
特异度	$Specificity = \frac{TN}{TN + FP}$	实际值为 Negative 的全部值中，算法判断正确的比例
F1 分数	$F1\ Score = \frac{2Precision * Recall}{Precision + Recall}$	F1-Score 值综合了 Precision 与 Recall 的结果。F1-Score 的值介于 0 到 1 之间，1 代表模型的最佳输出，0 代表模型的最差输出。
AUC 值	$AUC = \frac{(1 + TP - FP)}{2}$	AUC 值越接近于 1，表示分类效果越好。

表 6-3 真实值和预测值的评价指标

评价指标	公式	意义
均方误差 (Mean Square Error, MSE)	$MSE = \frac{1}{n} \sum_{j=1}^n I_{i,j} (\hat{y}_{i,j} - y_{i,j})^2$	MSE 直接反映真实值和预测值之间数据的离散情况。取值范围为 $[0, +\infty]$ ，当真实值和预测值完全一致时，MSE 为 0。MSE 越小，预测的误差越低。
均方根误差 (Root Mean Square Error, RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n I_{i,j} (\hat{y}_{i,j} - y_{i,j})^2}$	RMSE 是 MSE 的开根号结果，数量级上可以更直观显示误差的结果。取值范围为 $[0, +\infty]$ 。RMSE 越小，预测的误差越低。
平均绝对误差 (Mean Absolute Error)	$MAE = \frac{1}{n} \sum_{j=1}^n I_{i,j}  \hat{y}_{i,j} - y_{i,j} $	MAE 反映数据真实值和预测值之间差的平均。MAE 具有实际意义，可以反映实际数据的误差大小。MAE 的取值范围为 $[0, +\infty]$ ，它的值越小，表明预测的误差越小。

关于社区安全感的预测结果与真实结果比较，本章节通过相关评价指标来验证。常用的评价指标如表 6-3。

### 6.2.3 变量设定

#### 6.2.3.1 个体安全感知预测模型的输入变量解释结果变量

本章节用于个体安全感知预测模型建模的结果变量为基于问卷数据的居民不安全感指标，原始不安全感指标由被抢风险感知、被偷风险感知等五项指标的均值组成（与第四、五章节一致）。由于本章节研究任务是建模预测，情感感知类变量本身具有难以预测的特定，因此本研究将不安全感进行二值化操作（得分 $\leq 3$  记为 0，表示有安全感；得分 $> 3$  记为 1，表示无安全感），以期获取更高的预测精度，服务于后期社区安全感模拟任务。

#### 6.2.3.2 个体安全感知预测模型的输入变量

虽然第四章节的线性回归模型分析表明只有部分区域环境变量对安全感知有显著作用，但是并不是说明其他变量就与居民安全感知完全无关。一些因素与安全感知相关性强，但是没有显著的因果联系。一些因素，在线性回归模型结果中表现为不显著，但是在其他回归模型中的结果显著。因此，本研究根据理论分析和实际需要，除去那些共线性高的变量和不能用于模拟的变量，保留其他变量。最终的输入变量和之前第四章线性回归模型的变量基本一致。

本章节提出了个体人口统计特征、社区结构特征和社区微环境指标等共三方面的若干指标来预测个体的居民安全感，具体见表 6-4 和图 6-4。

### 6.2.4 研究单元

前述章节的研究区为广州市 7 个中心城区，共计 1874 个人口普查数据。本研究会结合街景微环境数据和手机数据进行社区安全感知的空间模拟，由于开放数据的不均匀特征，少量社区没有抓取到街景数据。同时，少量社区没有对应的手机数据指标。因此，除开少量没有街景数据的社区和没有手机数据覆盖的社区，这两类社区一共构成了缺失值社区，缺失值社区共有 121 个。因此，本研究最后模拟出的有社区安全感知值的社区共有 1753 个。本章节用到的调研社区也比之前第四、五章节的 82 个社区少 2 个，为 80 个。

具体来看，本研究在个体安全感知预测建模的过程，由于会用到问卷数据的

安全感知作为结果变量,因此相关的调研社区为80个,个体问卷样本为1701个。在运用个体安全感知模型进行社区安全感知模拟的时候,用到的相关社区为1753个。

表 6-4 个体安全感知模型的输入变量

尺度	类型	数据	指标解释	数据源	理论基础	
个体因素	人口统计特征	年龄	个体年龄, 定序变量, 从1到4, 不同年龄段。	问卷数据	脆弱性模型	
		性别	个体性别, 0 为男性, 1 为女性			
		受教育程度	个体的受教育水平, 定序变量, 从 1 到 7			
		户口	个体户口, 0 为外地户口, 1 为本地户口			
社区因素	社区结构指标	受教育比例	社区大专及以上人口占总人口比例	手机数据/ 统计数据	环境失序 理论、 环境设计预 防犯罪理论	
		外来人口比例	社区非广州户口人口占总人口比例			
		人口密度	社区总人口与土地面积的比例 (人/每平方千米)	手机数据/ 统计数据		
		商业设施比例	社区商业设施数占总 POI 数目的比例			
		政府机构比例	社区政府机构数占总 POI 数目的比例	POI 设施点数据 POI 设施点数据		
		犯罪率	社区犯罪案件数占总人数的比例乘以 100000 (每十万人的犯罪数)			
		NDVI 指数	归一化植被指数			
微环境指标		路灯比例	社区有路灯街景图片占该社区总街景图片的比例	遥感产品数据 街景数据 街景数据 街景数据	环境失序 理论、 环境设计预 防犯罪理论	
		天空开阔指数	社区所有街景图片的天空占比值的均值			
		绿视率	社区所有街景图片的植被占比值的均值			
		垃圾度	社区有建筑垃圾图片占该社区所有街景图片的比例			

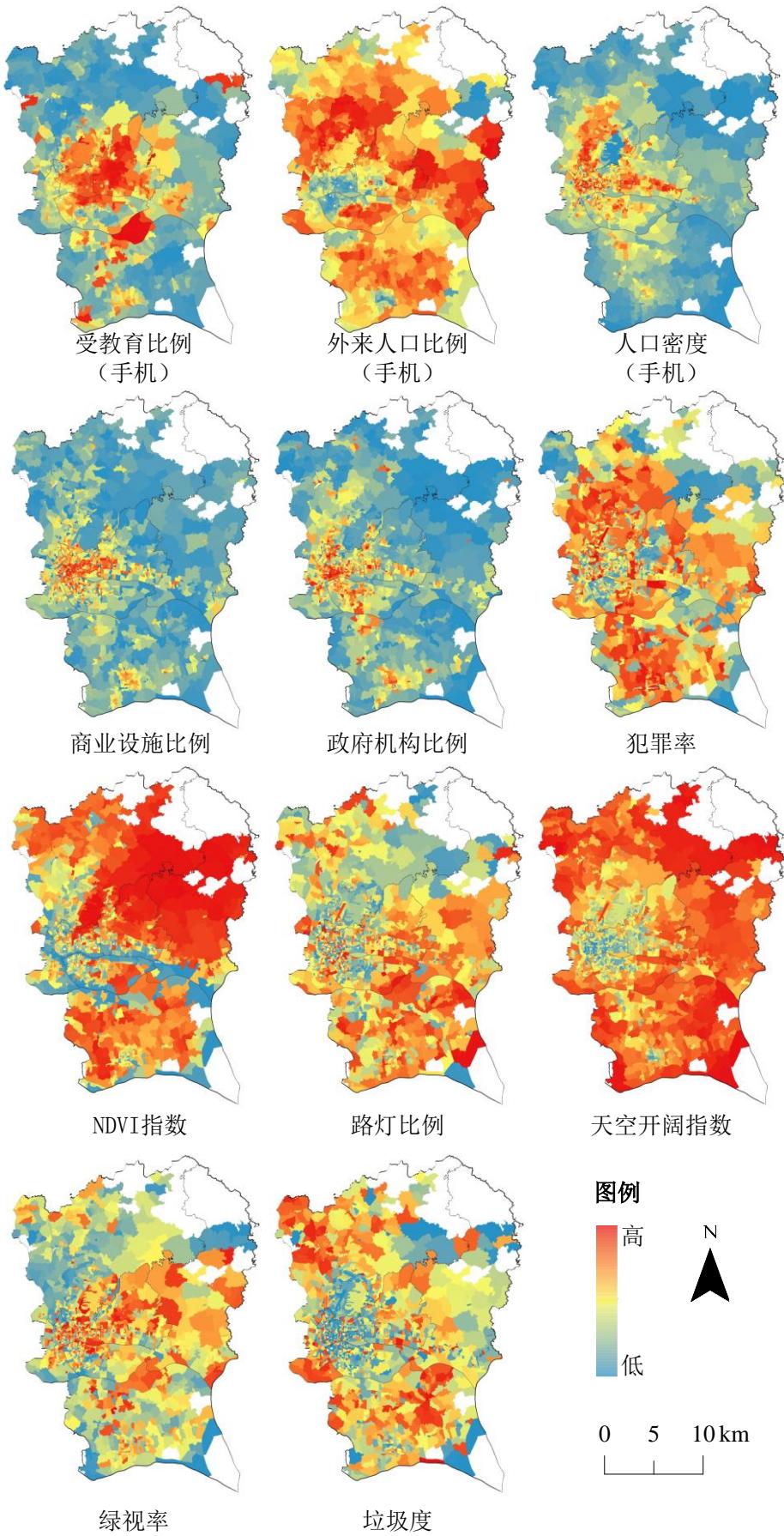


图 6-4 输入变量的社区环境指标

## 6.3 实验结果与分析

### 6.3.1 基于手机数据的个体安全感知预测模型的输入变量修正

#### 6.3.1.1 基于手机数据的社区统计指标确定

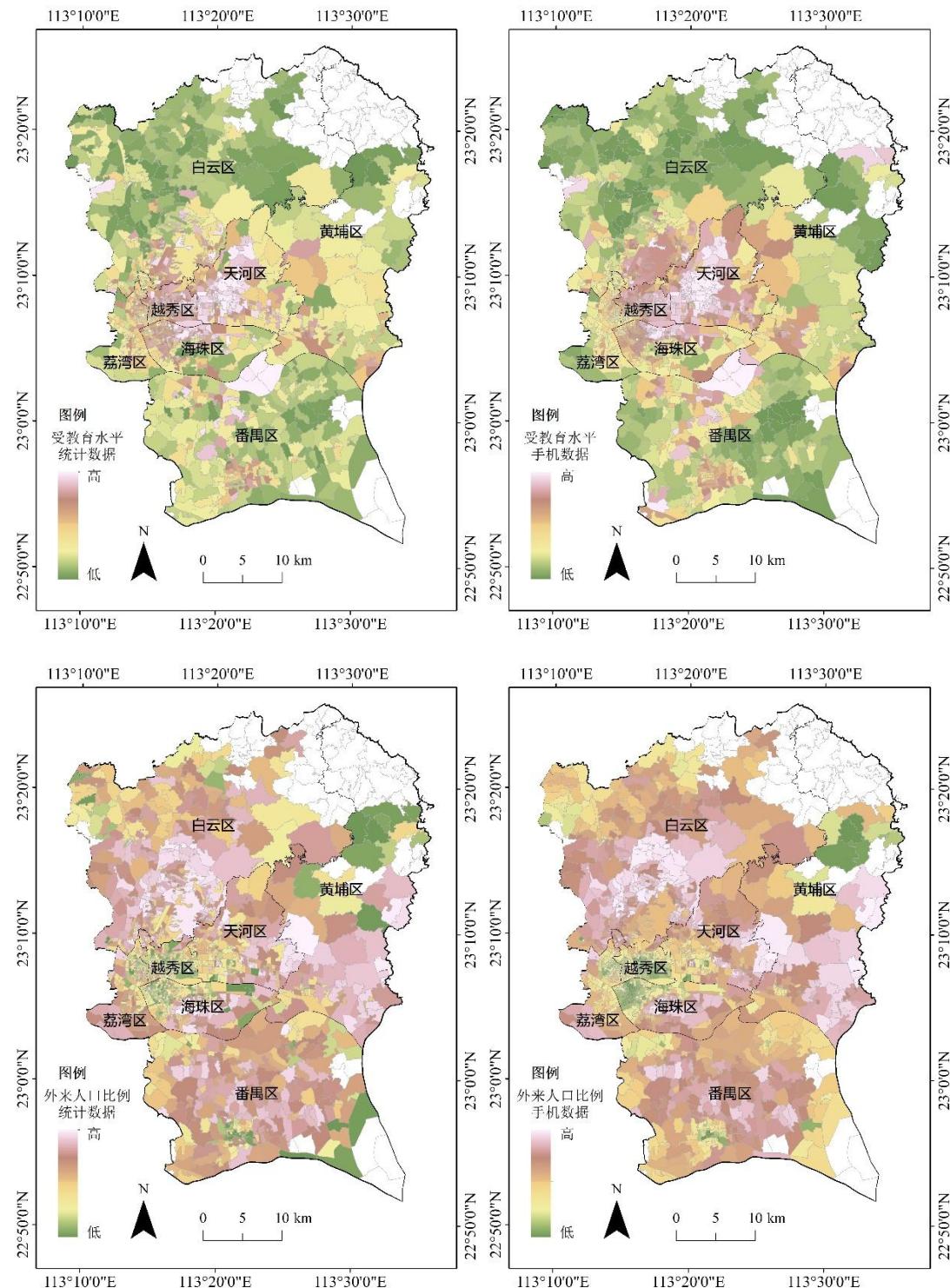


图 6-5 不同数据源的受教育水平和外来人口比例指标对比

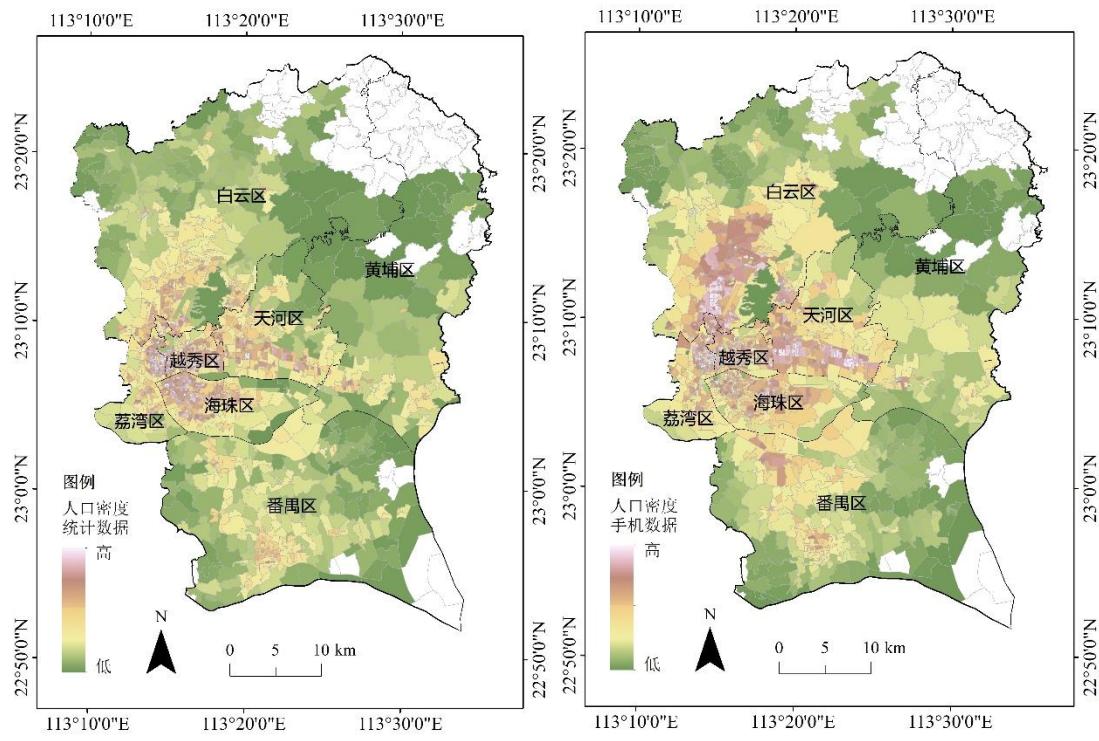


图 6-6 不同数据源的人口密度指标对比

本研究目的是通过大数据的手段获取时空粒度比较高的安全感知模型。因此，如果采用传统的人口普查统计数据，模型的时效性较低。因此，本文拟通过中国联通的手机数据，来代替传统的人口普查统计数据。

本研究拟替代的统计变量有 3 个，分别是社区外来人口比例、社区人口密度以及社区受教育水平。基于手机数据的社区统计变量的确定原理为，根据某时间段个体手机数据在某个社区停留的时间长度，从而确定某时间段该个体的所属社区。根据一个个体在 0 点至早上 5 点的停留地所在社区，从而确定该个体的居住社区，并匹配六普统计数据为该个体的居住社区数据。

在夜间 9 点至 12 点期间，识别落在该社区范围内的个体。人口密度的计算原理就是在这个时间段该社区范围手机的用户数量，再除以土地面积。受教育水平就是在这个时间段该社区范围内所有手机用户的居住地的受教育水平的平均值。外来人口比例就是在这个时间段该社区范围内手机用户的居住地的外来人口比例的平均值。经过处理，基于手机数据的社区统计指标如图 6-5 和图 6-6。

### 6.3.1.2 不同类型社区指标对模型解释力的对比

本研究拟用基于手机数据的社区统计指标代替基于统计数据的社区统计指

标。在进行变量直接替代前，本研究采用分别将基于手机数据的社区统计指标和基于统计数据的社区统计指标放入安全感知预测模型的方式，通过模型效果的评估，检验基于手机数据的社区统计指标是否比基于统计数据的社区统计指标更能准确预测安全感知。

由于本章节的安全感知是二分变量，本研究选取二元 logistics 回归模型，输入相关输入变量，进行模型测试。相关程序在 Stata/SE 15.1 软件中运行。模型结果如表 6-5 所示，模型 2 的 AIC 和 BIC 值都略大于模型 1 中对应的值。这表明，基于手机数据的社区指标是略优于基于统计数据的社区指标。在这个实验前提下，在接下来的步骤中，本研究直接将基于统计数据的社区指标替换为基于手机数据的社区指标。

表 6-5 基于二项式 logistics 回归的模型对比

	模型 1 (基于统计数据的社区指标)	模型 2 (基于手机数据的社区指标)
AIC	1913.88	1905.666
BIC	2001.276	1993.062

### 6.3.2 机器学习的数据预处理

根据2016广州市社区安全问卷及其对应的社区环境数据集，本研究共收集了1741个个体的15种属性。在进行模型训练之前，本文基于皮尔逊相关系数进行相关分析，所有预测变量的相关性低于0.5，变量之间不存在共线性问题。数据集中没有缺失值，因此不需要处理缺失值的操作。

本章节结果变量安全感知，无安全感的人群明显小于有安全感人群(0.243:0.757)，因此安全感知预测是一个非平衡分类问题，需要进行数据重采样，使无安全感和有安全感人群的大致一致，以保证模型训练的可靠性和稳定性。SMOTE (Synthetic Minority Oversampling Technique) 方法通过对少数类实现过采样，可以实现样本类别的平衡 (Chawla et al., 2002)，在保证数据质量的同时提升模型预测效果。因此，在训练数据集中<sup>6</sup>，本文通过SMOTE方法对少数类样本

<sup>6</sup> 经过试验，对总体样本进行 SMOTE 过采样，然后划分训练集和测试集，得到的预测效果更好，随机森林模型的总体精度可达 80.12%。但是，在实际对安全感知进行预测时，安全感知的两类也是不平衡分布。因此，本研究最后决定，还是先划分训练集和测试集，再只对训练集进行 SMOTE 过采样。

进行过采样操作，最终两类样本比例达到1:1。

安全感知的两类样本不平衡，相比于直接从数据集中按比例抽取划分训练集和测试集，根据调研问卷中有安全感、无安全感人群的比重再按比例抽取划分训练和测试集更为科学。因此，本研究分别对调研问卷有安全感、无安全感人群按比例随机抽取70%数据，再汇总作为模型训练集，剩余30%数据作为测试集。对模型训练集进行SMOTE操作，保证训练集中的正负样本比例达到平衡。

机器学习的相关代码操作和命令使用R语言完成。

### 6.3.3. 基于二项式 logistics 回归模型的安全感知预测结果

#### 6.3.3.1 模型构建及参数调整

根据之前的模型结构，本部分假设二项式 logistics 回归模型如下，同时根据之前章节中对训练样本数据的处理，建立居民安全感知的预测模型，分析各输入变量和安全感知的关系。具体的模型如下：

$$P = P(y = 1 | X_1, X_2, \dots, X_{15}) = \frac{e^{\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{15} X_{15} + \mu}}{e^{\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{15} X_{15} + \mu} + 1} \quad (6-9)$$

根据以上参数设置，利用 R 软件中的二项式 logistics 模型（family=“binomial”），得到各个回归系数（表 6-6）。

表 6-6 二项式 logistics 回归模型估计

变量	名称	Estimate	Std. Error	z-value	Pr(> z )	
	(Intercept)	1.892	0.674	2.809	0.005	**
X1	年龄	-0.073	0.078	-0.938	0.348	
X2	性别	0.579	0.126	4.611	0.000	***
X3	户口状态	-0.168	0.140	-1.203	0.229	
X3	受教育程度	0.056	0.062	0.898	0.369	
X4	社区街道路灯比例	0.809	0.639	1.265	0.206	
X5	社区街景绿视率	-2.575	1.361	-1.892	0.059	.
X6	社区街景垃圾度	-3.533	1.427	-2.475	0.013	*
X7	社区街景天空开阔程度	-3.071	1.443	-2.129	0.033	*
X8	社区商业设施比例	0.001	0.000	3.308	0.001	***
X9	社区政府机构比例	-0.002	0.001	-3.319	0.001	***
X10	NDVI	-0.586	0.850	-0.689	0.491	
X11	犯罪率	0.000	0.000	-3.355	0.001	***
X12	手机数据社区受教育程度	-1.213	0.471	-2.575	0.010	*
X13	手机数据社区外来人口比	-0.027	0.398	-0.069	0.945	
X14	手机数据社区人口密度	-0.001	0.000	-4.724	0.000	***

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1

AIC: 1574.8

结果发现，相当部分变量的系数不显著。因此，使用逐步回归法（step 函数）对变量进行进一步的选择。新的结果显示（表 6-7），所有变量的  $P$  值小于 0.1，通过了显著性检验。虽然模型的各变量达到显著性检验要求，但是要保证整个模型是有意义的，整个模型也需要通过统计检验。因此，对模型进行了卡方检验。

表 6-7 二项式 logistics 回归模型逐步回归

	Estimate	Std. Error	z-value	Pr(> z )	
(Intercept)	2.250	0.510	4.413	0.000	***
年龄	-0.119	0.066	-1.810	0.070	.
性别	0.554	0.124	4.467	0.000	***
社区街道路灯比例	0.918	0.616	1.492	0.136	
社区街景绿视率	-2.928	1.136	-2.578	0.010	**
街景垃圾度	-3.380	1.406	-2.404	0.016	*
社区街景天空开阔程度	-3.517	1.200	-2.930	0.003	**
社区商业设施比例	0.001	0.000	3.325	0.001	***
社区政府机构比例	-0.002	0.001	-3.342	0.001	***
犯罪率	0.000	0.000	-3.168	0.002	**
手机数据社区受教育程度	-1.111	0.454	-2.447	0.014	*
手机数据社区人口密度	-0.001	0.000	-4.749	0.000	***

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

AIC: 1569

卡方检验的结果显示（表 6-8），随着变量逐步添加到模型，模型最终通过了显著性检验。因此，由以上变量组成的二项式 logistics 回归模型是有意义的。

表 6-8 卡方检验结果

	Df	Deviance	Resid.Df	Resid.Dev	Pr(>Chi)	
NULL			1183	1641.2		
年龄	1.000	0.569	1182.000	1640.700	0.451	
性别	1.000	21.146	1181.000	1619.500	0.000	***
社区街道路灯比例	1.000	1.595	1180.000	1617.900	0.207	
社区街景绿视率	1.000	7.441	1179.000	1610.500	0.006	**
社区街景垃圾度	1.000	3.432	1178.000	1607.100	0.064	.
社区街景天空开阔程度	1.000	1.185	1177.000	1605.900	0.276	
社区商业设施比例	1.000	6.452	1176.000	1599.400	0.011	*
社区政府机构比例	1.000	8.283	1175.000	1591.200	0.004	**
犯罪率	1.000	10.254	1174.000	1580.900	0.001	**
手机数据社区受教育程度	1.000	10.342	1173.000	1570.500	0.001	**
手机数据社区人口密度	1.000	25.530	1172.000	1545.000	0.000	***

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

### 6.3.3.2 模型精度分析

接下来，通过混淆矩阵计算模型对测试集的预测精度。同时，为了更方便地衡量模型的优劣，本文通过 ROC 曲线的覆盖率来可视化表示模型的整体预测能力。

表 6-9 二项式 logistic 回归模型混淆矩阵结果

混淆矩阵	预测值		总数	分类精度
	有安全感	无安全感		
实际值	有安全感	206	190	396 52.02%
	无安全感	49	78	127 61.42%
	总体精度	-	-	523 54.3%

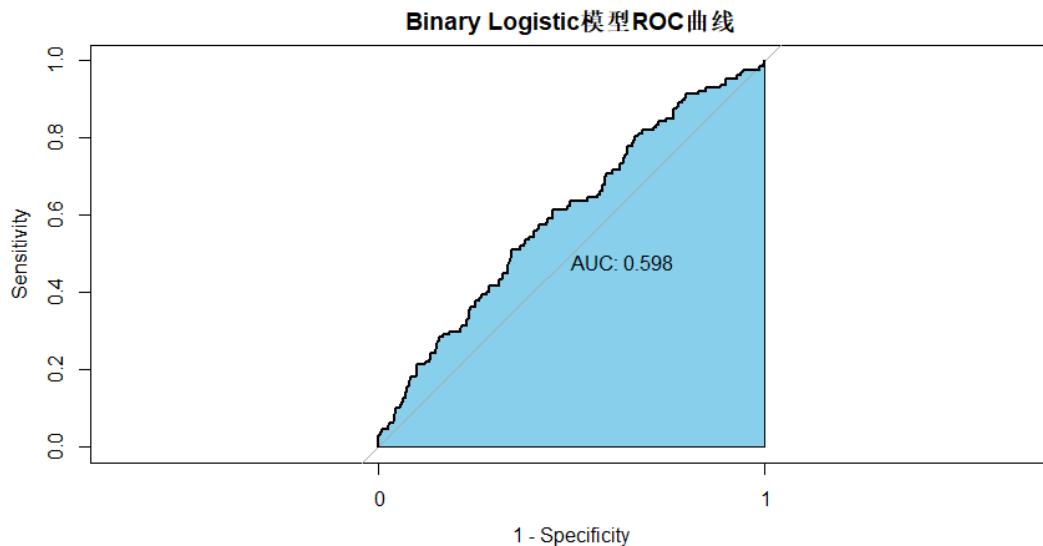


图 6-7 二项式 logistics 模型 ROC 曲线结果

由表 6-9 可知，模型对无安全感人群的预测能力高于有安全感人群的预测，但是预测结果都不太理想，总体精度在 54.3%。因此二项式 logistics 模型对于个体安全感的预测能力不强（特别是对有安全感的人群）。

二值分类器还可以用 ROC (Receiver Operating Characteristic) 曲线和 AUC 值来评价模型的效果。如图 6-7，图中 ROC 曲线的 AUC 值为 0.598，表明该模型的预测效果一般。

### 6.3.4. 基于 RF 随机森林模型的安全感知预测结果

#### 6.3.4.1 模型构建及参数调整

本文基于 R “randomForest” 程序包实现随机森林算法，利用袋外数据 (out-of-bag, OOB) 无偏估计来获取多种参数设置下模型的 OOB 误差值变化情况。要保证最优的分类精度，有必要进行调参操作。随机森林算法的两个重要参数分别为 ntree 和 mtry。其中，ntree 是随机森林包含的决策树数目，R 中的默认值是 500。mtry

为用于二叉树的指标数目,对于分类模型,通常是模型所有变量的平方根(叶雷,2017)。

由于本研究中的数据集变量为15个,模型问题是分类问题,所以本文选择3和4为可能的mtry值。在此基础上,ntree的确定一般是通过试算选取最优结果。本研究通过设置2~10逐2变化、10~100逐30变化,和100~1500逐100变化的实验,测试了不同决策树数目的结果。所得结果如图6-8所示。

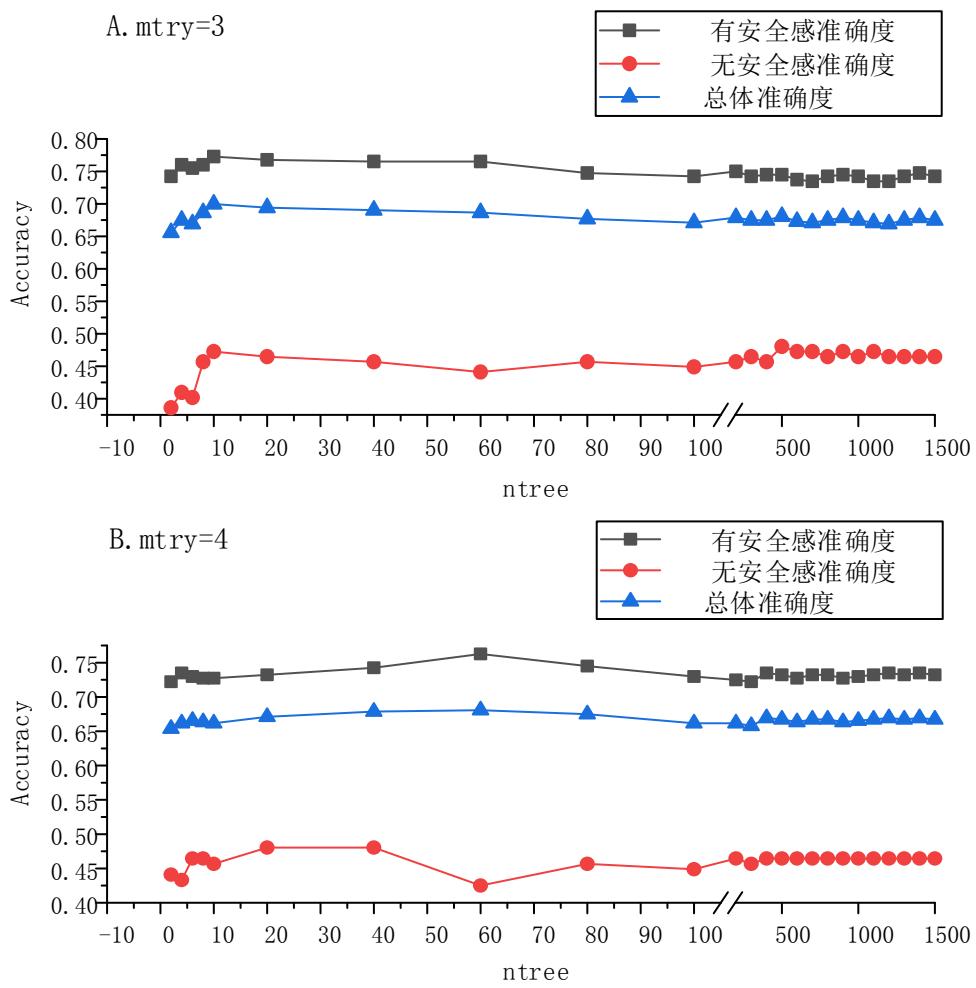


图 6-8 不同树选择下的分类精度

从图6-8可以看出,mtry=3时的预测准确率整体上大于mtry=4时的情况。对于mtry=3时,当树的个数在0~10区间变化时,总体上各类精度随着树的个数增加而显著提高。当树的个数在10~100区间变化时,精度随树的个数增加先滑落再提高。当树的个数在100~1500区间变化时,精度随树的个数增加而逐渐升高后保持稳定。当mtry=3,ntree=10时,各类精度值都达到最佳。因此,根据不同mtry和ntree的调参实验结果,考虑到模型精度评价的Accuracy指标值,本研究的随机森林模型最终确定的参数为mtry=3,ntree=10。

### 6.3.4.2 模型精度分析

使用上述最优参数，利用70%的经过SMOTE方法重采样后的数据作为训练样本，对随机森林模型进行训练。剩余30%的数据作为测试样本。把训练好的模型运用到测试集。表6-10的混淆矩阵给出了本文模型的总体分类精度和对两类安全感的分类精度。可知，总体分类精度为69.98%，对有安全感和无安全感人群的分类精度分别为77.27%和47.24%。该方法对有安全感人群的分类精度最高，对无安全感人群的分类精度相对较低。

表 6-10 混淆矩阵

混淆矩阵		预测值		总数	分类精度
		有安全感	无安全感		
实际值	有安全感	306	90	396	77.27%
	无安全感	67	60	127	47.24%
	总体精度	-	-	523	69.98%

随机森林可以在高维输入向量中执行特征的选取，因此，即使噪声数据，它的效果也较好。本研究将之前收集到的所有指标作为输入向量，随机地给各个特征变量加入噪声干扰。随机森林根据模型准确率的浮动，使用准确率降低的程度测度特征重要性排序。根据OOB误差得到的平均精度减少量如图6-9A所示，该值越大说明所对应的变量越重要。

图 6-9B 为通过随机森林模型得到的各个特征重要性的排序（基于 GINI 系数）。15 个特征 F1-F15 排序分别为：F1 受教育程度；F2 年龄；F3 个体所在社区手机数据受教育水平；F4 个体所在社区手机数据人口密度；F5 性别；F6 个体所在社区街景绿视率；F7 个体所在社区犯罪率；F8 个体所在社区路灯比例；F9 个体所在社区街景天空开阔程度；F10 个体所在社区政府机构比例；F11 个体所在社区商业设施比例；F12 个体所在社区手机数据外来人口比例；F13 个体所在社区遥感植被指数 NDVI；F14 户口状态；F15 个体所在社区街景垃圾度。

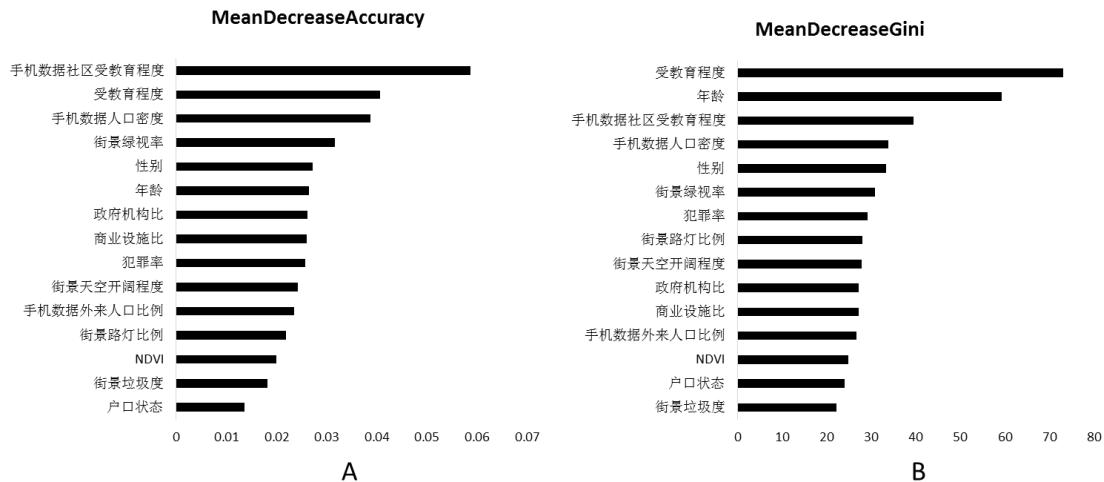


图 6-9 重要性程度排序

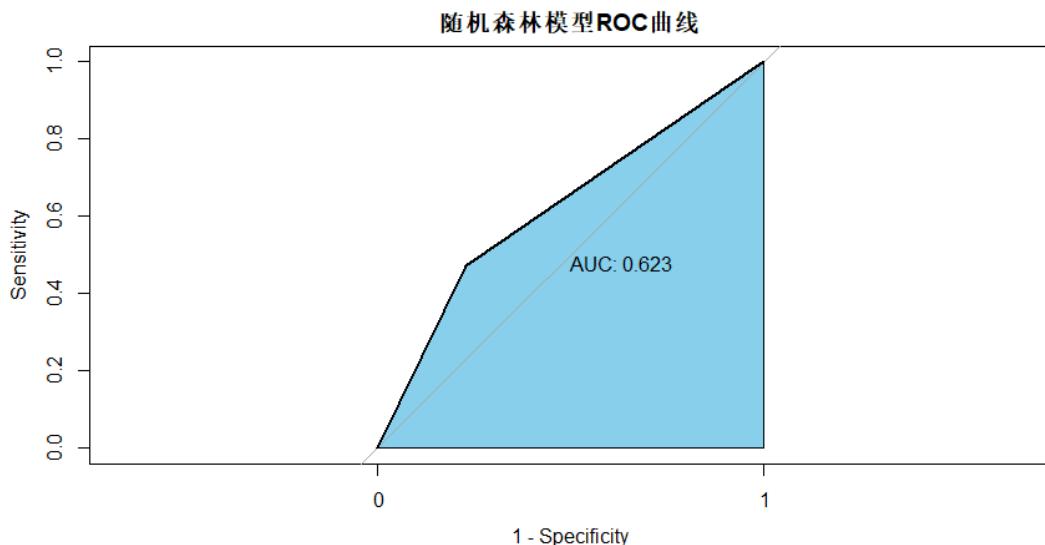


图 6-10 随机森林模型 ROC 曲线结果

如图 6-10, ROC 曲线的结果显示, AUC 值为 0.623, 模型的整体拟合效果处于可以接受的区间。

### 6.3.5 基于 SVM 支持向量机模型的安全感知预测结果

安全感知 SVM 预测模型的构建过程如下。第一步, 选取用于安全感知预测的指标和数据集。第二步, 完成数据预处理。第三步, 调整支持向量机, 选取最优惩罚参数 cost 和核函数参数 gamma。第四步, 利用训练好的 SVM 进行预测。第五步, 根据预测的结果选择最佳方案, 建立安全感知 SVM 预测模型。实验研究中, 本研究利用 R 语言及其“e1071”程序包完成支持向量机模型的运行。

### 6.3.5.1 模型构建及参数调整

输入模型的特征变量与二项式 logistics 模型和随机森林中的预测变量一致。确定输入特征后，数据预处理相关输入变量。特别是在 SVM 模型中，一个经验性的结论是标准化后的数据的模型预测结果更优。因此，本研究对输入特征进行数据 0-1 标准化处理，然后再进行训练和测试数据集。

SVM 模型有若干的参数需要确定，本文使用默认的参数完成相关设置。其中，对于核函数的选择，本文选择默认的“radial”核函数（径向基核函数，也称为高斯核函数）。对于成本函数和 gamma 函数，本文设置成本函数为默认的值 1，gamma 函数设置为数据维度的倒数（1 / 数据维度）。SVM-Type，即本模型的类别设置为 C 分类器模型（C-classification）。在此设置下，完成初步的模型运算。

表 6-11 初步 SVM 混淆矩阵结果

混淆矩阵		预测值		总数	分类精度
		有安全感	无安全感		
实际值	有安全感	217	46	263	82.51%
	无安全感	179	81	260	31.15%
	总体精度	-	-	523	56.98%

表 6-11 混淆矩阵显示，总体的分类精度是 56.98%，可见真实值和预测值的一致性程度一般，对于有安全感人群的分类精度大幅高于无安全感人群的分类精度。有必要进行进一步的模型调整。

模型调整的第一步，通过调用 R 语言中的 *tune.svm* 命令调整 SVM。该命令是采用试错法，通过测试不同的 cost 和 gamma 值，从而得到不同的预测效果，从而得到最佳的 gamma 和惩罚因子。

本研究将惩罚因子 cost 分别设置为 cost=10 和 cost=100，gamma 参数范围设置为 gama=10 ^(-6:-1)，共有 12 组的 SVM 参数组合。通过 10 次交叉试验法计算每次组合的误差偏差，确定误差最小的参数组合。*Tune.svm* 的表现结果如表 6-12。可知，在 gama=0.01、cost= 100 的模型中，算法的效果最好，error=0.2475。当得到最佳参数后，使用由 *tuning* 函数得到的最佳参数设置再训练一个新的居民安全感知预测的支持向量机。

表 6-12 SVM 模型调参汇总结果

	<b>gamma</b>	<b>cost</b>	<b>error</b>	<b>dispersion</b>
1	1.00E-06	10	0.5143284	0.03243808
2	1.00E-05	10	0.5143284	0.03243808
3	1.00E-04	10	0.4434696	0.03278056
4	1.00E-03	10	0.4130679	0.04385451
5	1.00E-02	10	0.3767554	0.03252338
6	1.00E-01	10	0.2601054	0.0347374
7	1.00E-06	100	0.5143284	0.03243808
8	1.00E-05	100	0.4451574	0.03147342
9	1.00E-04	100	0.4214998	0.04125108
10	1.00E-03	100	0.3978778	0.04957031
11	1.00E-02	100	0.3235152	0.03139443
12	1.00E-01	100	0.2475004	0.04132486

### 6.3.5.2 模型精度分析

参数配置完成后，调用进行居民安全感知预测。整理测试集的真实分类和预测分类，然后获取相关系数，得到混淆矩阵，结合 ROC 曲线的参数，评价最后的模型效果。

表 6-13 最优 SVM 模型混淆矩阵结果

	<b>混淆矩阵</b>	<b>预测值</b>		<b>总数</b>	<b>分类精度</b>
		<b>有安全感</b>	<b>无安全感</b>		
实际值	有安全感	250	59	309	80.91%
	无安全感	146	68	214	31.78%
	总体精度	-	-	523	60.8%

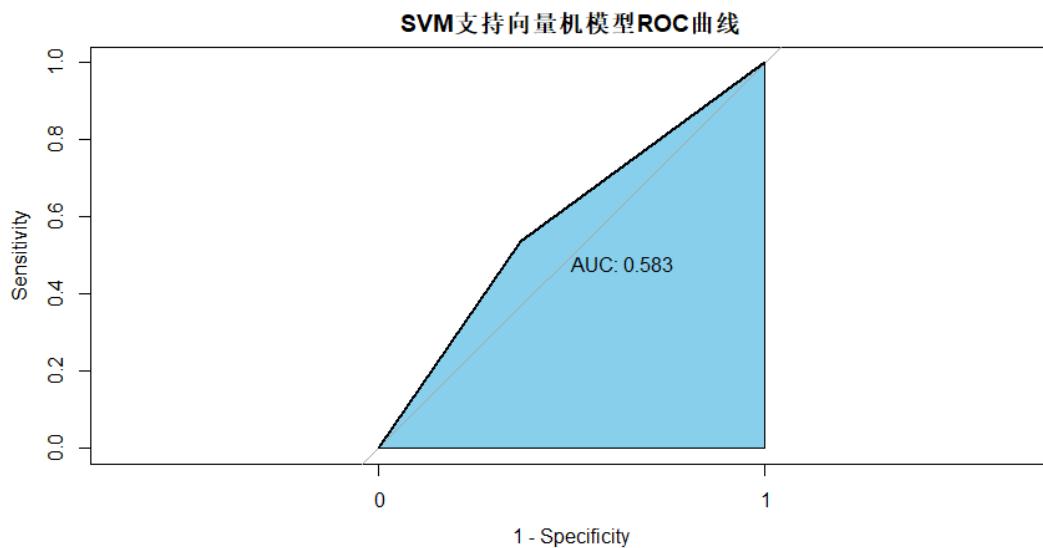


图 6-11 支持向量机模型 ROC 曲线结果

通过图 6-11 和表 6-13 可知，模型总体的预测精度达到 60.8%，比参数优化前的模型效果提升了接近 4 个百分点，说明模型调参的效果较为显著。这个预测精度比传统的 logistics 回归模型的精度高，但是仍然比随机森林的模型精度低。

此外，对无安全感人群的分类效果还是较差。

### 6.3.6 安全感知分类的各算法对比分析

以上，分别实现了二项式 logistics 回归算法、随机森林算法和支持向量机算法。为了比较各模型的效果，由于本章节中安全感知是一个二分变量，预测任务本质上是一个分类模型，所以本研究通过混淆矩阵构建一系列评价指标来综合评判（表 6-14）。

表 6-14 各算法不同分类安全感知评价精度指标

算法	Accuracy	Precision	Recall	Specificity	F1-score
二元 Logistics 回归算法	0.543	0.614	0.291	0.808	0.395
支持向量机算法（SVM）	0.608	0.318	0.535	0.631	0.399
随机森林算法(RF)	0.700	0.472	0.4	0.82	0.433

由表 6-14 可以得出，综合 Accuracy、Precision、Recall、Specificity 以及 F1 分数的比较，随机森林算法比广义线性 logistics 回归和支持向量机算法的预测结果好，也从一个方面说明了应用随机森林进行安全感知评估的先进性。从总体效果看，随机森林模型在安全感建模上具有应用前景。因为 logistics 回归模型的决策边界是线性的，因此 logistics 回归是广义线性模型。而在非线性的条件下，具有这种方法的规律还原性较差。因此，整体回归效果劣于随机森林算法。随机森林模型通过自学习的手段进行数据分析和非线性拟合，既确保了训练数据集拟合的精准度，也确保在样本之外泛化预测时，具有较好的扩展能力。比较随机森林模型和支持向量机模型的结果，说明对于安全感知这种不平衡类别主观情感感知的预测，随机森林模型的分类效果更好。

在应用中，随机森林模型也有一些不足。当使用随机森林模型进行安全感预测时，不能获得具体的函数表达式，也不能直观获得每个特征因素对安全感影响的大小。而传统的多元线性回归或 logistics 回归可以完成这些任务。

## 6.4 社区安全感知预测与验证

本研究的主要目的是服务于实际的政府社会治理工作。安全感知是一种情感状态，对个体安全感知预测的现实意义有限。但是个体安全感知的空间化对政府

可以产生实际的指导作用。如果一个社区具有较低的安全感状态，而周围社区的安全感状态较高，即使该社区的客观犯罪率不高，管理机构也应该采取相关手段提升该社区的居民安全感水平。因此，在训练好的模型的基础上，本文拟用该模型模拟研究区内 9328963 个常住人口的安全感知得分，进而聚合到社区尺度，完成广州市社区尺度安全感知的空间分布。

### 6.4.1 研究区所有个体属性数据生成及预测输入数据确定

个体安全感知模型的预测变量共有 15 个。其中，14 个变量为社区尺度因素，可以直接用社区数据赋予到研究区内所有个体上。4 个变量为个体变量（性别、年龄、个体受教育程度和户口状态）。由于不可能直接获取到研究区内所有个体的具体个人属性特征，同时考虑到本研究的目的在于最终聚合到社区尺度，计算出社区尺度的安全感知，因此本研究按照“六普”（第六次人口普查数据）中研究区社区的性别、年龄、受教育程度、户口指标等字段的人数数据特征，生成个体模拟数据。

根据六普数据中以社区为统计单元的各项指标，本文编写一个批量生成数据的 Python 程序算法，最终生成了 9328963 条个体属性数据。生成数据的要求如下：第一，每个社区的各项指标符合人口统计数据的分布。比如，A 社区有 3000 人口，其中，1600 人口为男性，1400 人口为女性。那么生成的数据的性别这一列中，男女的比例和数目均与统计数据相匹配。第二，生成的数据需要是随机分布于各行。即，不能前 1600 行的年龄字段是男性，后 1400 人口的年龄字段是女性，这样会导致模型结果产生较大的误差。正确的分布应该为这 3000 行的年龄字段中，男性和女性是随机分布于各行。具体示例如表 6-15 & 16。

表 6-15 模拟数据分布表

<b>community</b>	<b>case</b>	<b>age</b>	<b>hukou</b>	<b>edu</b>	<b>gen</b>
440111103208	1	2	1	4	0
440111103208	2	1	1	3	0
440111103208	3	4	1	4	0
440111103208	4	3	0	3	0
440111103208	5	3	1	3	1
440111103208	6	1	1	3	0
440111103208	7	3	1	4	1
440111103208	8	2	1	3	1
440111103208	9	3	1	3	0
440111103208	10	4	1	2	1
440111103208	11	1	1	2	1
440111103208	12	3	1	4	1
...	...	...	...	...	...
440116003001	9328955	1	0	3	1
440116003001	9328956	2	0	4	0
440116003001	9328957	2	1	2	1
440116003001	9328958	2	0	4	0
440116003001	9328959	1	0	3	0
440116003001	9328960	1	0	4	1
440116003001	9328961	1	0	5	0
440116003001	9328962	1	0	4	0
440116003001	9328963	1	0	5	1

表 6-16 模拟数据分布表指标解释

列名	名称	解释
community	社区	社区代码
case	编号	个体代码编号
age	年龄	根据六普数据中每个社区内不同年龄段的人数分布，生成对应样本的个体年龄分布。年龄分段，从 1 到 7。
hukou	户口	根据六普数据每个社区内不同户口状态的人数分布，生成对应样本的户口属性。0 为非广州户口，1 为广州户口。
edu	受教育程度	根据六普数据每个社区内不同受教育程度的人数分布，生成对应样本的个体受教育程度分布。0 为男性，1 为女性。六普数据中有街道尺度，但无社区尺度的性别数据。因此，根据六普数据每个街道内性别占比，作为对应街道内所属所有社区的性别比例。然后，生成对应社区内的个体性别分布。0 为男性，1 为女性。
gen	性别	

在完成全部模拟后,共模拟得到有 4 个变量属性的 932 万余条个体属性数据。然后,把已有的 11 个社区环境变量赋值到对应的个体上(基于个体的社区单元)。最后,得到九百余万条个体数据,包含 15 个特征变量,构成  $9328963 \times 15$  的矩阵。

#### 6.4.2 研究区个体安全感知值预测与验证

利用已经训练好的基于随机森林算法的安全感知预测模型,本文对全部 932 万余条个体数据的安全感知进行了预测(得分为 0 表明有安全感,得分为 1 表明无安全感)。然后,基于社区单元,本文以平均值的方法聚合得到各社区的安全感得分,分数区间在  $[0,1]$ 。

进一步,本研究对预测效果进行检验。对于真实值和预测值的比较,常用的评价指标如表 6-17。

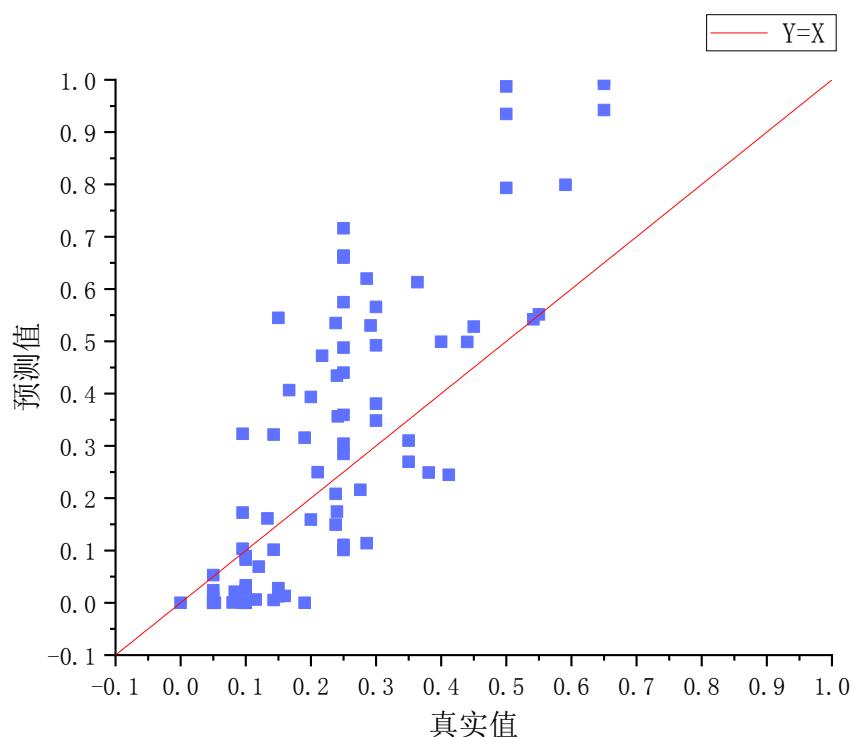


图 6-52 真实值与预测值散点图比较

表 6-17 基于样本社区的验证结果指标

统计指标	相关系数	MSE	RMSE	MAE
值	0.826	0.037	0.191	0.34

本研究将 80 个样本社区<sup>7</sup>中每个社区所有受访者的二值化安全感知以均值方式聚合到社区尺度，得到 80 个调研社区的真实社区安全感知。挑选出模拟得到的 80 个对应社区的安全感知值与真实值进行对比，检验本文安全感知空间模拟的效果。基于一系列评价指标的模型效果检验结果见表 7-18 和附表 2。

从图 6-12 可知，所有样本社区的预测值和真实值都靠近  $Y=X$  趋势线附近，说明验证效果较好。不过，整体上预测值相对于真实值稍微偏高。

从表 7-17 可知，验证结果指标 MSE、RMSE、MAE 的值都较小，相关系数值较高，表明相关社区的真实值和预测值的吻合程度较高，模型的验证效果较好。本研究的基于小样本问卷调查实现大范围的空间安全感知预测具有较高的预测精度。

#### 6.4.3 社区安全感知空间化结果

如图 6-13，得到研究区社区不安全感的分布图。从图 6-13 和表 7-18 可知，荔湾区、海珠区和白云区是不安全感最高的区域，而天河区、黄埔区、番禺区是安全感最高的区域。这说明安全感知的空间分布并不与犯罪率的分布以及经济发展水平密切吻合，这也与之前的文献结论一致 (Scarborough et al., 2010; Jing et al., 2020)。天河区属于广州新的经济中心，年轻人口多，经济水平高，城市景观有序，可能导致了不安全感的低水平状态。荔湾区属于老城区，老年人口占比大，人口密度大，街道空间品质还有待提高，可能导致了不安全感的高水平状态。具体来看，可以发现一些安全感知分布的局部特征。从图 7-15 可知，一般街道绿化水平较高的区域，社区不安全感较低。比如，番禺大学城、二沙岛、中山大学、花城广场等区域，以及外围城市化水平较低的郊区区域。而一些社会环境比较混杂的区域的不安全感较高，比如白云机场北部、海珠区南部布匹市场区域。

表 6-18 行政区尺度安全感知预测结果

统计指标	海珠区	番禺区	黄埔区	荔湾区	天河区	越秀区	白云区
预测值	社区不安全感	0.661	0.501	0.508	0.665	0.413	0.558
	排序	2	6	5	1	7	3

<sup>7</sup>由于手机数据和本文研究区范围略有差异，因此除开没有街景数据的社区，本文还有少数没有手机数据覆盖的社区，故本章节用到的样本社区也比之前第四、五章节的 82 个社区少 2 个，为 80 个样本社区。

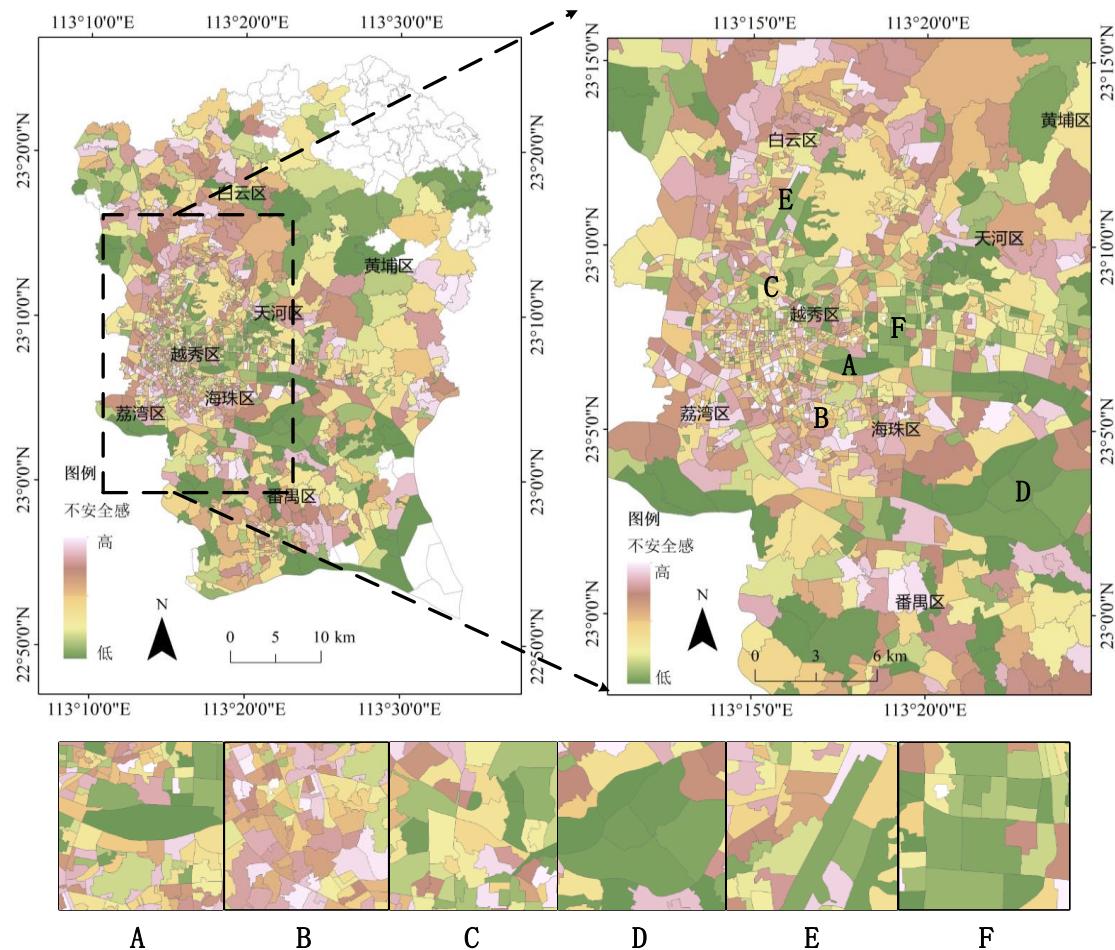


图 6-63 广州市研究区不安全感空间分布

(A 是二沙岛附近; B 是中大南门布匹市场附近; C 是广州火车站附近; D 是番禺大学城附近; E 是白云国际机场附近; F 是花城广场附近)

## 6.5 小结

本章节在前述章节的基础上，利用线性和非线性回归方法，以问卷数据的安全感知及其影响因素为样本，训练得到了多种安全感知预测模型。最后，运用效果最优的随机森林预测模型对研究区内所有个体的安全感知进行了实际预测，并对安全感知的社区级空间分布进行了可视化。主要结论如下：

(1) 本章节实现了基于多源数据的安全感知预测，并通过传统 logistics 线性回归模型，以及支持向量机和随机森林两种非线性机器学习模型的对比，发现非线性的随机森林模型对于个体安全感知的预测能力表现最好，综合预测精度达到 0.7。该发现为今后的安全感知感模型建模提供了新的思路，在不是特别关注模型中影响因素的影响程度的情况下，可以由传统的线性模型转为非线性模型，

以提升模型的解释效果。同时，在研究不同影响因素是否是影响安全感知的问题上，非线性模型可以通过模型解释度的变化来证明不同影响因素的效果。

(2) 本章节实现了从个体安全感知预测到社区安全感知预测的一套有效分析方法。其中关键点如下，第一，本研究以社区安全感知预测为最终目的，对研究区范围类所有个体以社区统计数据为基础进行数据模拟，模拟得到百万级的个体属性数据。第二，通过基于问卷数据的个体安全感知进行训练建模，最后通过样本区域个体安全感知聚和的社区安全感知进行结果验证（本文的验证精度MSE为0.037），以此保证社区安全感知空间模拟的科学性。综合来看，本文这种方式的处理，既弥补了社区安全感知预测时个体数据缺乏的问题，又保证了模拟的科学性。这一套由小数据到大数据的模拟方法可以实现安全感知的实时、高效、科学的预测。在实际工作中，通过小样本的问卷发放，利用大样本的多源大数据，可以实现城市区域精细化时空尺度的安全感知刻画。

(3) 本章节的安全感知空间模拟结果显示，安全感知空间分布与社区经济水平、犯罪率水平并不是直接相关，一些不安全感知较高的区域分布在中心城区。这一方面验证了之前文献中关于安全感知与犯罪率之间的关系(Wyant, 2008; Alkimim, Clarke, & Oliveira, 2013)，另一方面也为实际的城市安全感提升工作提供了科学依据。通过空间可视化以及进一步的空间格局分析，可以发现社区安全感知的高值和低值区域，相关部门进而可以进行实际的治安防控工作，提升城市居民的安全感。

## 第七章 总结与展望

### 7.1 研究结论

安全感知是人们经常会有的感受。安全感知和区域犯罪率并不总是保持一致 (Scarborough et al., 2010; Jing et al., 2020)。一些犯罪率高的地区，居民的被害风险感知低。一些犯罪率低的地区，居民的被害风险感知高。因此，社会治安防控工作，不仅要着手降低犯罪率，还要特别关注如何提升治安安全感知。安全感知的研究包括安全感知的测量、影响因素分析、空间化模拟等方面。近年来，安全感知的测量从传统静态的问卷测量到使用移动设备实现动态测量逐渐转变。影响因素分析从传统单独关注个体因素、社会环境因素、自然环境因素，到关注个体、社会环境、自然环境交互作用的分析。空间安全感知测量从传统基于问卷数据的空间化到使用移动设备实现空间化。

本研究的最终目的是基于小样本的问卷调查，实现安全感知的空间模拟。这一套空间模拟方法的实现，可以以小成本的方式快速计算出整个城市区域不同空间分辨率下的安全感知空间分布，从而识别出安全感知的高值和低值区域，科学指导实际的社会治安防控工作。

从城市管理和社会防控角度出发，安全感知的空间模拟不仅需要达到高的空间分辨率，还需要达到高的时间分辨率。从实现高时空分辨率的安全感知空间模拟出发，在考虑安全感知的关键影响因素时，需要选取时空分辨率高的指标。街景图像是实时获取的数据源，基于路网分布，包含丰富的城市环境信息，真实地刻画了城市居民的日常环境暴露。从日常活动理论出发，基于街景图像的环境信息是影响居民个体安全感知的重要因素。因此，本研究尝试使用街景图像作为居民安全感空间模拟的重要数据源，提取一系列街景微环境指标，进而分析街景微环境指标与安全感知的关系和影响机制。在此基础上，运用多源数据，建立个体安全感预测模型。最后，运用训练好的个体安全感预测模型，同时借助科学模拟得出的广州市千万级别个体人口数据，实现广州市所有个体居民的安全感知预测。得出模拟结果后，通过均值化的方式实现安全感知的空间模拟和空间制图。

综合来看，本研究得出了如下结论：

### (1) 微环境因素的加入有利于增强安全感知模型的解释能力

比起由传统影响因素组成的模型，加入微环境因素，安全感知预测模型的拟合效果增强，说明微环境因素对安全感知具有重要作用。显著性分析也发现，一些微环境因素显著地影响安全感知。具体来看，绿视率是显著影响安全感知的因素。由于街景植被的可视率较高，为居民的日常交流提供了公共场所，而且街景植被对居民的身心健康有益处，可以有效降低居民的攻击性。因此，区域高水平的街景绿化可以提升居民的安全感知。

### (2) 街景微环境因素对安全感知不仅有直接作用，还存在间接作用。

在研究街景微环境对安全感知的直接作用基础上，本研究继续探讨了街景微环境的间接作用。根据客观环境-主观环境-安全感知的理论机制，本研究深入分析了特定街景微环境（分别为街景绿视率和街景路灯比例）是如何通过主观环境作用于安全感知。研究结论表明，① 客观微环境对于主观安全感知的影响机制受感知的环境因素中介。② “客观环境-感知环境-主观安全感知”的作用机制中既存在中介作用，也存在调节作用和交互作用。这说明，安全感知的主客观影响因素之间存在相互作用，安全感知的影响机制较为复杂。在进一步的安全感知预测时，有必要综合比较线性和非线性的建模方法。

### (3) 对于个体安全感知的预测，非线性模型比传统线性模型的预测能力更好。

由于安全感知空间模拟的需要，本研究需要对安全感知进行预测。为获取较好的预测效果，本研究分别对比了线性模型和非线性模型，具体包括传统Logistics回归分析模型、支持向量机模型、随机森林模型。并通过划分训练集和测试集进行预测准确度的比较。结果显示，相比于线性回归模型，两种机器学习模型的预测效果更优。而相比于支持向量机模型，随机森林在本研究中的个体安全感知预测模型的预测效果更好。因此，本研究验证了对于个安全感知的预测任务，非线性模型比线性模型更优。相应地，本研究提出了一种基于随机森林的个体安全感知预测方法。

### (4) 本研究实现了一种基于小样本的问卷调查开展大规模的社区安全感知空间模拟的方法

本研究的最终目的是进行安全感知的空间模拟。基于小样本的问卷调查，本研究开发了一套安全感知空间模拟方法。主要的步骤为，基于街景图像获取街景

微环境指标，再检验街景微环境指标对安全感知的直接和间接作用。在此基础上，得到个体安全感知的预测模型，同时模拟出广州市所有个体的基本属性，从而进行所有个体的安全感知预测。最后，基于个体安全感知的均值实现社区安全感知空间化，进行效果验证，同时达到空间模拟和空间制图的目的。

### 7.2 论文创新点

本文有以下几点创新：

(1) 通过加入微环境因素，设计了更加完善的居民安全感知预测模型

本研究从环境暴露的角度，提取了基于街景图像的微环境指标，把这些指标加入到安全感知模型。模型结果发现，微环境因素是影响安全感知的重要因素，而且模型的拟合效果得到提高。因此，在传统安全感知预测模型的基础上，本研究提出了一个更加完善的安全感知模型。

(2) 揭示了客观微环境因素-主观感知因素-主观安全感知之间的影响机制

在传统环境与安全感知关系研究的基础上，本研究提出并验证了客观微环境-主观感知环境-主观安全感知之间的作用机制。本研究表明，街景微环境对于安全感知的作用是通过感知建成环境失序中介，且受到直接受害经历调节和交互。

(3) 验证了非线性模型（随机森林模型）对个体安全感知预测的良好效果

安全感知作为一种情感感知，影响机制复杂。在分析微环境对安全感知的直接作用和间接作用基础上，本研究发现安全感知的影响机制复杂，影响因素多样。因此，本研究通过比较两种非线性建模方式和一种线性回归建模方式，发现了非线性方法可以更加准确地预测居民的安全感知。其中，随机森林模型的预测效果最优，预测准确度较线性模型提升了 15.68%。

(4) 开发了一套基于问卷调查实现基于大数据的居民安全感知空间模拟的方法

情感感知的高时空分辨率模拟一直是难以突破的科学问题。本研究运用高时空分辨率的影响因子，模拟出广州市所有居民与安全感知相联系的属性数据，选取最优的安全感知预测模型，完成了个体安全感知的预测，实现了一整套安全感知空间模拟。本研究的方法可以实现基于小样本数据完成大尺度安全感知空间模拟，对于实际的城市规划和城市管理工作具有重要实际作用。

### 7.3 不足之处与展望

本文的研究也存在一些不足之处，主要有以下几点：

(1) 街景微环境指标的利用程度。本研究中在检验环境与安全感知的关系，以及进行居民安全感知的空间模拟的研究中，将基于街景的微环境指标聚合在社区尺度的区域上。这种方式的处理是由于本研究虽然调查了受访者个体的信息，但是没有精确的地理坐标信息，只能确定个体的所在社区位置，所以将环境指标也聚合在社区尺度上。然而，从更真实地衡量个体日常活动的地理背景角度出发，可以再将街景微环境的范围缩小，进行分析基于日常活动暴露的街景微环境对个体居民安全感知的影响。这样，可以分析出不同时间空间尺度下微环境对个体安全感的作用。

此外，一些环境指标，聚合到社区尺度，可能会影响该指标对于安全感知的真实作用机制。比如，犯罪率指标是聚合在社区尺度上，但是实际犯罪活动发生在少数热点区域。这样，即使是在高犯罪率的社区，实际上大多数地点是安全的。在今后的研究中，需要进一步考虑相关因素的处理。

(2) 街景微环境指标的局限性。首先，街景图像数据在时间和空间分辨率上都存在局限性，本研究获取的街景图像数据和问卷数据的时间段不是直接重合。街景数据不是每年更新，而且由于开放数据平台的技术限制，街景图片的获取也只能获取最新一期的街景，而不能获取历史特定时间段的街景。此外，一些人们经常行走的小路和巷道，街景车却不能进入，因此街景图像也不能完全均衡反应个体的日常环境暴露，它的空间分辨率在一些区域上不是均衡的。

(3) 安全感知影响因素还需要进一步考虑。安全感知作为一种主观感受，受到许多不同类别因素的影响。比如，白天和夜间的安全感知受到的环境影响因素就可能不一致。而且，一些街景微环境因素还可以细化，比如，绿视率可以进一步细化为不同植被健康程度的街景指标。如果能够获取真实的社区尺度的收入水平或城市12345投诉热线的城市环境指标数据，可以更加真实和多角度刻画社区尺度的环境指标。

(4) 调节效应和中介效应分析可进一步优化。本研究的机制分析是将所有变量视为显变量。然而，区分显变量和潜变量可能会得到一些更有意义的结果。比如，安全感知和非正式社会融入这类型变量，本身不能实际直接被测量，而只

能通过一些题项来间接测量，这些类型的变量可能更适合用于潜变量来表示。而绿视率等是客观存在的变量，可以当作显变量。因此，在接下来的工作中，可尝试运用结构方程模型，针对具体问题，研究特定环境要素对安全感知的影响机制。此外，本研究只是分析了绿视率这一反应环境失序程度的街景微环境指标，没有同时分析其他主客观环境因素与安全感知之间的机制。进一步的研究工作中，可以同时考虑更多因素的中介效应分析（如链式中介效应）和调节效应分析（如跨层调节效应分析），以探测出更多的影响路径。

(5)本研究没有考虑地理背景不确定性问题(Kwan, 2012)。一些研究发现，个体的不安全感知不止受个体所属社区的影响，而且会受到周围社区环境的影响(Rountree, 1998; Breetzke & Pearson, 2014; Boessen & Hipp, 2015)。个体特别是女性返回家的过程中，如果在自身社区范围，可能不安全感知不高，但是如果处于所属社区范围外的社区环境，(由于对环境陌生)则不安全感知可能较高。因此，邻近社区的微环境对居民安全感知也可能存在影响。



## 参考文献

- [1] Abbott, J., & McGrath, S. A. (2017). The effect of victimization severity on perceived risk of victimization: Analyses using an international sample. *Victims & Offenders*, 12(4), 587-609.
- [2] Adams, R. E., & Serpe, R. T. (2000). Social integration, fear of crime, and life satisfaction. *Sociological perspectives*, 43(4), 605-629.
- [3] Aguiléra, V., Allio, S., Benezech, V., Combes, F., & Milion, C. (2014). Using cell phone data to measure quality of service and passenger flows of Paris transit system. *Transportation Research Part C: Emerging Technologies*, 43, 198-211.
- [4] Akaike, H. (1987). Factor analysis and AIC. In *Selected papers of hirotugu akaike* (pp. 371-386). Springer, New York, NY.
- [5] Alda, E., Bennett, R. R., & Morabito, M. S. (2017). Confidence in the police and the fear of crime in the developing world. *Policing: An International Journal of Police Strategies & Management*.
- [6] Alfaro-Beracoechea, L., Puente, A., Da Costa, S., Ruvalcaba, N., & Páez, D. (2018). Effects of fear of crime on subjective well-being: A meta-analytic review. *European Journal of Psychology Applied to Legal Context*, 10(2), 89-96.
- [7] Alkimim, A., Clarke, K. C., & Oliveira, F. S. (2013). Fear, crime, and space: The case of Viçosa, Brazil. *Applied Geography*, 42, 124-132.
- [8] Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health*, 68(6), 578-583.
- [9] Badland, H. M., Opit, S., Witten, K., Kearns, R. A., & Mavoa, S. (2010). Can virtual streetscape audits reliably replace physical streetscape audits?. *Journal of Urban Health*, 87(6), 1007-1016.
- [10] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12), 2481-2495.
- [11] Baran, P. K., Tabrizian, P., Zhai, Y., Smith, J. W., & Floyd, M. F. (2018). An exploratory study of perceived safety in a neighborhood park using immersive virtual environments. *Urban Forestry & Urban Greening*, 35, 72-81.
- [12] Barni, D., Vieno, A., Roccato, M., & Russo, S. (2016). Basic personal values, the country's crime rate and the fear of crime. *Social Indicators Research*, 129(3), 1057-1074.
- [13] Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6), 1173.
- [14] Bentham, J. (1983). *The Collected Works of Jeremy Bentham: Deontology. Together with a Table of the Springs of Action and the Article on Utilitarianism*.
- [15] Boessen, A., & Hipp, J. R. (2015). Close - ups and the scale of ecology: Land uses and the geography of social context and crime. *Criminology*, 53(3), 399-426.
- [16] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152).
- [17] Box, S., Hale, C., & Andrews, G. (1988). Explaining fear of crime. *The British Journal of Criminology*, 28(3), 340-356.

- [18] Breetzke, G. D., & Pearson, A. L. (2014). The fear factor: Examining the spatial variability of recorded crime on the fear of crime. *Applied Geography*, 46, 45-52.
- [19] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123-140.
- [20] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [21] Britto, S., Stoddart, D., & Ugwu, J. (2018). Perceptually contemporaneous offenses: Gender and fear of crime among African-American university students. *Journal of Ethnicity in Criminal Justice*, 16(2), 117-136.
- [22] Brunson, L., Kuo, F. E., & Sullivan, W. C. (2001). Resident appropriation of defensible space in public housing: Implications for safety and community. *Environment and Behavior*, 33(5), 626-652.
- [23] Brunton - Smith, I., & Sturgis, P. (2011). Do neighborhoods generate fear of crime? An empirical test using the British Crime Survey. *Criminology*, 49(2), 331-369.
- [24] Brunton-Smith, I., Jackson, J., & Sutherland, A. (2014). Bridging structure and perception: On the neighbourhood ecology of beliefs and worries about violent crime. *British Journal of Criminology*, 54(4), 503-526.
- [25] Castro-Toledo, F. J., Perea-Garc á, J. O., Bautista-Ortu ño, R., & Mitkidis, P. (2017). Influence of environmental variables on fear of crime: Comparing self-report data with physiological measures in an experimental design. *Journal of Experimental Criminology*, 13(4), 537-545.
- [26] Chadee, D., Ng Ying, N. K., Chadee, M., & Heath, L. (2019). Fear of crime: the influence of general fear, risk, and time perspective. *Journal of interpersonal violence*, 34(6), 1224-1246.
- [27] Chan, K. W., & Zhang, L. (1999). The hukou system and rural-urban migration in China: Processes and changes. *The China Quarterly*, (160), 818-855.
- [28] Charreire, H., Mackenbach, J. D., Ouasti, M., Lakerveld, J., Compernolle, S., Ben-Rebah, M., ... & Oppert, J. M. (2014). Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the SPOTLIGHT project). *Health & place*, 25, 1-9.
- [29] Chataway, M. L., Hart, T. C., Coomber, R., & Bond, C. (2017). The geography of crime fear: A pilot study exploring event-based perceptions of risk using mobile technology. *Applied geography*, 86, 300-307.
- [30] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.
- [31] Cho, J. T., & Park, J. (2017). Exploring the effects of CCTV upon fear of crime: A multi-level approach in Seoul. *International Journal of Law, Crime and Justice*, 49, 35-45.
- [32] Cho, Y., Jeong, H., Choi, A., & Sung, M. (2019). Design of a connected security lighting system for pedestrian safety in smart cities. *Sustainability*, 11(5), 1308.
- [33] Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American sociological review*, 588-608.
- [34] Conklin, J. E. (1971). Dimensions of community response to the crime problem. *Social problems*, 18(3), 373-385.
- [35] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- [36] Cozens, P. M., Neale, R. H., Whitaker, J., Hillier, D., & Graham, M. (2003). A critical review of street lighting, crime and fear of crime in the British city. *Crime Prevention and Community Safety*, 5(2), 7-24.
- [37] Cutler, A., & Zhao, G. (2001). Pert-perfect random tree ensembles. *Computing Science and Statistics*, 33, 490-497.

- [38] Doran, B. J., & Burgess, M. B. (2012). Why is fear of crime a serious social problem?. In *Putting Fear of Crime on the Map* (pp. 9-23). Springer, New York, NY.
- [39] Du, Y., & Law, J. (2016). How do vegetation density and transportation network density affect crime across an urban central-peripheral gradient? A case study in Kitchener—Waterloo, Ontario. *ISPRS International Journal of Geo-Information*, 5(7), 118.
- [40] Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016, October). Deep learning the city: Quantifying urban perception at a global scale. In European conference on computer vision (pp. 196-212). Springer, Cham.
- [41] Ellaway, A., Macintyre, S., & Bonnefoy, X. (2005). Graffiti, greenery, and obesity in adults: secondary analysis of European cross sectional survey. *Bmj*, 331(7517), 611-612.
- [42] Espinosa, J. E., Velastin, S. A., & Branch, J. W. (2017, November). Vehicle detection using alex net and faster R-CNN deep learning models: a comparative study. In *International Visual Informatics Conference* (pp. 3-15). Springer, Cham.
- [43] Farrall, S., & Gadd, D. (2004). Research note: The frequency of the fear of crime. *British Journal of Criminology*, 44(1), 127-132.
- [44] Farrall, S., Bannister, J., Ditton, J., & Gilchrist, E. (1997). Questioning the measurement of the ‘fear of crime’: Findings from a major methodological study. *The British Journal of Criminology*, 37(4), 658-679.
- [45] Ferraro, K. & LaGrange, R. (1987). The measurement of fear of crime. *Sociological Inquiry*, 57(1), 70-97.
- [46] Foster, S., Giles-Corti, B., & Knuiman, M. (2010). Neighbourhood design and fear of crime: A social-ecological examination of the correlates of residents’ fear in new suburban housing developments. *Health & place*, 16(6), 1156-1165.
- [47] Foster, S., Knuiman, M., Wood, L., & Giles-Corti, B. (2013). Suburban neighbourhood design: Associations with fear of crime versus perceived crime risk. *Journal of Environmental Psychology*, 36, 112-117.
- [48] Foster, S., Villanueva, K., Wood, L., Christian, H., & Giles-Corti, B. (2014). The impact of parents’ fear of strangers and perceptions of informal social control on children’s independent mobility. *Health & place*, 26, 60-68.
- [49] Franklin, T. W., Franklin, C. A., & Fearn, N. E. (2008). A multilevel analysis of the vulnerability, disorder, and social integration models of fear of crime. *Social Justice Research*, 21(2), 204-227.
- [50] Gabriel, U., & Greve, W. (2003). The psychology of fear of crime. Conceptual and methodological perspectives. *British Journal of Criminology*, 43(3), 600-614.
- [51] Garofalo, J. (1981). The fear of crime: Causes and consequences. *J. Crim. L. & Criminology*, 72, 839.
- [52] Gates, L. B., & Rohe, W. M. (1987). Fear and reactions to crime: A revised model. *Urban Affairs Quarterly*, 22(3), 425-453.
- [53] Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences*, 201700035.
- [54] Gibson, C. L., Zhao, J., Lovrich, N. P., & Gaffney, M. J. (2002). Social integration, individual perceptions of collective efficacy, and fear of crime in three cities. *Justice quarterly*, 19(3), 537-564.
- [55] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

- [56] Gong, F. Y., Zeng, Z. C., Zhang, F., Li, X., Ng, E., & Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, 134, 155-167.
- [57] Gray, E., Jackson, J., & Farrall, S. (2008). Reassessing the fear of crime. *European Journal of Criminology*, 5(3), 363-380.
- [58] Gray, E., Jackson, J., & Farrall, S. (2011). Feelings and functions in the fear of crime: Applying a new approach to victimisation insecurity. *The British journal of criminology*, 51(1), 75-94.
- [59] Guedes, I., Fernandes, P., & Cardoso, C. (2014). Studying the Contextual Cues Associated With Fear of Crime Through Eye Tracking Techniques. In ET4S@ GIScience (pp. 42-46).
- [60] Hale, C. (1996). Fear of crime: A review of the literature. *International review of Victimology*, 4(2), 79-150.
- [61] Hanslmaier, M. (2013). Crime, fear and subjective well-being: How victimization and street crime affect fear and life satisfaction. *European Journal of Criminology*, 10(5), 515-533.
- [62] Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication monographs*, 85(1), 4-40.
- [63] He, L., Páez, A., Jiao, J., An, P., Lu, C., Mao, W., & Long, D. (2020). Ambient Population and Larceny-Theft: A Spatial Analysis Using Mobile Phone Data. *ISPRS International Journal of Geo-Information*, 9(6), 342.
- [64] Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., & Wang, R. (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.
- [65] Helfgott, J. B., Parkin, W. S., Fisher, C., & Diaz, A. (2020). Misdemeanor arrests and community perceptions of fear of crime in Seattle. *Journal of Criminal Justice*, 69, 101695.
- [66] Herzog, T. R., & Bryce, A. G. (2007). Mystery and preference in within-forest settings. *Environment and Behavior*, 39(6), 779-796.
- [67] Ho, T. K. (1995, August). Random decision forests. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 1, pp. 278-282). IEEE.
- [68] Hodgkinson, T.; Gately, N.; McCue, J.; Shuhad, A.; Corrado, R.R.; Andresen, M.A. Fear of crime in an island paradise: Examining the generalizability of key theoretical constructs in the Maldivian context. *Int. Crim. Justice Rev.* 2017, 27, 108–125
- [69] Hough, M. (1995). Anxiety about crime: findings from the 1994 British Crime Survey.
- [70] Hunter, A., & Baumer, T. L. (1982). Street traffic, social integration, and fear of crime. *Sociological Inquiry*, 52(2), 122-131.
- [71] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7132-7141).
- [72] Jackson, J. (2004). Experience and expression: Social and cultural significance in the fear of crime. *British journal of criminology*, 44(6), 946-966.
- [73] Jackson, J. (2015). Cognitive closure and risk sensitivity in the fear of crime. *Legal and Criminological Psychology*, 20(2), 222-240.
- [74] Jackson, J., & Gouseti, I. (2012). Fear of crime and the psychology of risk. Jackson, J. and Gouseti, I., 'Fear of Crime and the Psychology of Risk', entry in Bruinsma,

- G. and Wesiburd, D.(eds.) Encyclopedia of Criminology and Criminal Justice, Springer-Verlag, Forthcoming.
- [75] Jacobs, J. (1992). The death and life of great American cities. 1961. New York: Vintage.
- [76] James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of applied psychology*, 69(2), 307.
- [77] Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790-794.
- [78] Jeffery, C. R. (1971). Crime prevention through environmental design (Vol. 91). Beverly Hills, CA: Sage publications.
- [79] Jiang, B., Deal, B., Pan, H., Larsen, L., Hsieh, C. H., Chang, C. Y., & Sullivan, W. C. (2017). Remotely-sensed imagery vs. eye-level photography: Evaluating associations among measurements of tree cover density. *Landscape and urban planning*, 157, 270-281.
- [80] Jing, F., Liu, L., Zhou, S., & Song, G. (2020). Examining the Relationship between Hukou Status, Perceived Neighborhood Conditions, and Fear of Crime in Guangzhou, China. *Sustainability*, 12(22), 9614.
- [81] Jing, F., Liu, L., Zhou, S., Song, J., Wang, L., Zhou, H., ... & Ma, R. (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.
- [82] Jorgensen, A., Hitchmough, J., & Calvert, T. (2002). Woodland spaces and edges: their impact on perception of safety and preference. *Landscape and urban planning*, 60(3), 135-150.
- [83] Kanan, J. W., & Pruitt, M. V. (2002). Modeling fear of crime and perceived victimization risk: The (in) significance of neighborhood integration. *Sociological inquiry*, 72(4), 527-548.
- [84] Karakus, O., McGarrell, E. F., & Basibuyuk, O. (2010). Fear of crime among citizens of Turkey. *Journal of Criminal Justice*, 38(2), 174-184.
- [85] Kaufman, Y. J., & Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE transactions on Geoscience and Remote Sensing*, 30(2), 261-270.
- [86] Kelly, C. M., Wilson, J. S., Baker, E. A., Miller, D. K., & Schootman, M. (2012). Using Google Street View to audit the built environment: inter-rater reliability results. *Annals of Behavioral Medicine*, 45(suppl\_1), S108-S112.
- [87] Kronqvist, K. (2014). Virtual observations of urban neighborhood physical disorder using Google street view. In The Stockholm Criminology Symposium, Stockholm, Sweden (2014) (pp. 169-169). The Swedish National Council for Crime Prevention (BRÅ).
- [88] Kuo, F. E., Bacaicoa, M., & Sullivan, W. C. (1998). Transforming inner-city landscapes: Trees, sense of safety, and preference. *Environment and behavior*, 30(1), 28-59.
- [89] Kupchik, A., & Bracy, N. L. (2009). The news media on school crime and violence: Constructing dangerousness and fueling fear. *Youth violence and juvenile justice*, 7(2), 136-155.
- [90] Kwan, M. P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102(5), 958-968.
- [91] LaGrange, R. L., & Ferraro, K. F. (1989). Assessing age and gender differences in perceived risk and fear of crime. *Criminology*, 27(4), 697-720.

- [92] Lee, J. S., Park, S., & Jung, S. (2016). Effect of crime prevention through environmental design (CPTED) measures on active living and fear of crime. *Sustainability*, 8(9), 872.
- [93] Lee, J., & Cho, S. (2018). The impact of crime rate, experience of crime, and fear of crime on residents' participation in association: studying 25 districts in the City of Seoul, South Korea. *Crime Prevention and Community Safety*, 20(3), 189-207.
- [94] Lee, S. W., Ellis, C. D., Kweon, B. S., & Hong, S. K. (2008). Relationship between landscape structure and neighborhood satisfaction in urbanized areas. *Landscape and Urban Planning*, 85(1), 60-70.
- [95] Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, 14(3), 675-685.
- [96] Lis, A., Pardela, L., & Iwankowski, P. (2019). Impact of vegetation on perceived safety and preference in city parks. *Sustainability*, 11(22), 6324.
- [97] Liska, A.E., Sanchirico, A. & Reed, M.D. (1988). Fear of Crime and Constrained Behavior: Specifying and Estimating a Reciprocal Effects Model. *Social Forces*, 66, 827-837.
- [98] Liu, J., Messner, S. F., Zhang, L., & Zhuo, Y. (2009). Socio-demographic correlates of fear of crime and the social context of contemporary urban China. *American journal of community psychology*, 44(1-2), 93-108.
- [99] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- [100] Maas, J., Spreeuwenberg, P., Van Winsum-Westra, M., Verheij, R. A., Vries, S., & Groenewegen, P. P. (2009). Is green space in the living environment associated with people's feelings of social safety?. *Environment and Planning A*, 41(7), 1763-1777.
- [101] Macassa, G., Wijersjö, R., Wijk, K., McGrath, C., Ahmadi, N., & Soares, J. (2017). Fear of crime and its relationship to self-reported health and stress among men. *Journal of public health research*, 6(3).
- [102] MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, 58, 593-614.
- [103] Maguire, E. R., Johnson, D., Kuhns, J. B., & Apostolos, R. (2017). The effects of community policing on fear of crime and perceived safety: findings from a pilot project in Trinidad and Tobago. *Policing and society*.
- [104] Mak, B. K., & Jim, C. Y. (2018). Examining fear-evoking factors in urban parks in Hong Kong. *Landscape and Urban Planning*, 171, 42-56.
- [105] Maruthaveeran, S. (2016). The perception of social safety in a green environment: a preliminary study at the Kepong Metropolitan Park. *Asian Journal of Environment-Behaviour Studies*, 1(1), 99-111.
- [106] May, D. C., Keith, S., Rader, N. E., & Dunaway, R. G. (2015). Predicting adolescent fear of crime through the lens of general strain theory. *Sociological Focus*, 48(2), 172-189.
- [107] May, D. C., Rader, N. E., & Goodrum, S. (2010). A gendered assessment of the "threat of victimization": Examining gender differences in fear of crime, perceived risk, avoidance, and defensive behaviors. *Criminal justice review*, 35(2), 159-182.
- [108] Mittal, G., Yagnik, K. B., Garg, M., & Krishnan, N. C. (2016, September). Spotgarbage: smartphone app to detect garbage using deep learning. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and*

- Ubiquitous Computing (pp. 940-945).
- [109] Mooney, S. J., Bader, M. D., Lovasi, G. S., Teitler, J. O., Koenen, K. C., Aiello, A. E., ... & Rundle, A. G. (2017). Street audits to measure neighborhood disorder: virtual or in-person?. *American journal of epidemiology*, 186(3), 265-273.
- [110] Mouratidis, K. (2019). The impact of urban tree cover on perceived safety. *Urban Forestry & Urban Greening*, 44, 126434.
- [111] Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of personality and social psychology*, 89(6), 852.
- [112] Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., & Hidalgo, C. A. (2015). Do people shape cities, or do cities shape people? The co-evolution of physical, social, and economic change in five major US cities (No. w21620). National Bureau of Economic Research.
- [113] Newman, O. (1972). Defensible space (p. 264). New York: Macmillan.
- [114] O'Brien, D. T., & Sampson, R. J. (2015). Public and private spheres of neighborhood disorder: Assessing pathways to violence using large-scale digital records. *Journal of research in crime and delinquency*, 52(4), 486-510.
- [115] O'brien, D. T., Sampson, R. J., & Winship, C. (2015). Econometrics in the age of big data: Measuring and assessing "broken windows" using large-scale administrative records. *Sociological Methodology*, 45(1), 101-147.
- [116] Ordonez, V., & Berg, T. L. (2014, September). Learning high-level judgments of urban perception. In European conference on computer vision (pp. 494-510). Springer, Cham.
- [117] Pain, R. (2000). Place, social relations and the fear of crime: a review. *Progress in human geography*, 24(3), 365-387.
- [118] Pánek, J., Pászto, V., & Marek, L. (2017). Mapping emotions: spatial distribution of safety perception in the city of Olomouc. In The rise of big spatial data (pp. 211-224). Springer, Cham.
- [119] Porter, J. R., Rader, N. E., & Cossman, J. S. (2012). Social disorganization and neighborhood fear: Examining the intersection of individual, community, and county characteristics. *American Journal of Criminal Justice*, 37(2), 229-245.
- [120] Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891.
- [121] Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, 42(1), 185-227.
- [122] Quinn, J. W., Mooney, S. J., Sheehan, D. M., Teitler, J. O., Neckerman, K. M., Kaufman, T. K., ... & Rundle, A. G. (2016). Neighborhood physical disorder in New York City. *Journal of maps*, 12(1), 53-60.
- [123] Rand, M. R. (2008). Criminal victimization, 2007 (NCJ224390). Washington, DC: US Department of Justice.
- [124] Reisig, M. D., & Parks, R. B. (2000). Experience, quality of life, and neighborhood context: A hierarchical analysis of satisfaction with police. *Justice quarterly*, 17(3), 607-630.
- [125] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*.
- [126] Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6), 1137-1149.

- [127] Richardson, E. A., & Mitchell, R. (2010). Gender differences in relationships between urban green space and health in the United Kingdom. *Social science & medicine*, 71(3), 568-575.
- [128] Roccato, M., Russo, S., & Vieno, A. (2011). Perceived community disorder moderates the relation between victimization and fear of crime. *Journal of Community Psychology*, 39(7), 884-888.
- [129] Roh, S., & Oliver, W. M. (2005). Effects of community policing upon fear of crime. *Policing: An International Journal of Police Strategies & Management*.
- [130] Rollwagen, H. (2016). The relationship between dwelling type and fear of crime. *Environment and Behavior*, 48(2), 365-387.
- [131] Ross, C. E., & Jang, S. J. (2000). Neighborhood disorder, fear, and mistrust: The buffering role of social ties with neighbors. *American journal of community psychology*, 28(4), 401-420.
- [132] Rountree, P. W. (1998). A reexamination of the crime-fear linkage. *Journal of research in crime and delinquency*, 35(3), 341-372.
- [133] Rühs, F., Greve, W., & Kappes, C. (2017). Coping with criminal victimization and fear of crime: The protective role of accommodative self - regulation. *Legal and criminological psychology*, 22(2), 359-377.
- [134] Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American journal of preventive medicine*, 40(1), 94-100.
- [135] Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. *PloS one*, 8(7), e68400.
- [136] Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of “broken windows”. *Social psychology quarterly*, 67(4), 319-342.
- [137] Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *science*, 277(5328), 918-924.
- [138] Scarborough, B. K., Like-Haislip, T. Z., Novak, K. J., Lucas, W. L., & Alarid, L. F. (2010). Assessing the relationship between individual characteristics, neighborhood context, and fear of crime. *Journal of Criminal Justice*, 38(4), 819-826.
- [139] Scheider, M. C., Rowell, T., & Bezdkian, V. (2003). The impact of citizen perceptions of community policing on fear of crime: Findings from twelve cities. *Police Quarterly*, 6(4), 363-386.
- [140] Schwartz, H., Eichstaedt, J., Kern, M., Dziurzynski, L., Lucas, R., Agrawal, M., ... & Ungar, L. (2013, June). Characterizing geographic variation in well-being using tweets. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 7, No. 1).
- [141] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of statistics*, 6(2), 461-464.
- [142] Shevky, E., & Williams, M. (1949). *The Social Area Analysis of Los Angeles: Analysis and Typology*.
- [143] Singer, A. J., Chouhy, C., Lehmann, P. S., Walzak, J. N., Gertz, M., & Biglin, S. (2019). Victimization, fear of crime, and trust in criminal justice institutions: A cross-national analysis. *Crime & Delinquency*, 65(6), 822-844.
- [144] Skinner, B. F. (1965). *Science and human behavior* (No. 92904). Simon and Schuster.
- [145] Snedker, Karen A. "Neighborhood conditions and fear of crime: A

- reconsideration of sex differences." *Crime & Delinquency* 61, no. 1 (2015): 45-70.
- [146] Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*, 13, 290-312.
- [147] Solymosi, R., Bowers, K. J., & Fujiyama, T. (2017). Crowdsourcing Subjective Perceptions of Neighbourhood Disorder: Interpreting Bias in Open Data. *The British Journal of Criminology*, 58(4), 944-967.
- [148] Solymosi, R., Bowers, K., & Fujiyama, T. (2015). Mapping fear of crime as a context - dependent everyday experience that varies in space and time. *Legal and Criminological Psychology*, 20(2), 193-211.
- [149] Solymosi, R., Buil-Gil, D., Vozmediano, L., & Guedes, I. S. (2020). Towards a place-based measure of fear of crime: A systematic review of app-based and crowdsourcing approaches. *Environment and Behavior*, 0013916520947114.
- [150] Song, G., Bernasco, W., Liu, L., Xiao, L., Zhou, S., & Liao, W. (2019). Crime feeds on legal activities: Daily mobility flows help to explain thieves' target location choices. *Journal of Quantitative Criminology*, 35(4), 831-854.
- [151] Stafford, M., Chandola, T., & Marmot, M. (2007). Association between fear of crime and mental health and physical functioning. *American journal of public health*, 97(11), 2076-2081.
- [152] Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS one*, 10(2), e0107042.
- [153] Stodolska, M., Shinew, K. J., Acevedo, J. C., & Roman, C. G. (2013). "I was born in the hood": Fear of crime, outdoor recreation and physical activity among Mexican-American urban adolescents. *Leisure Sciences*, 35(1), 1-15.
- [154] Tchinda, P. E., & Kim, S. N. (2020). The paradox of "eyes on the street": Pedestrian density and fear of crime in Yaoundé Cameroon. *Sustainability*, 12(13), 5300.
- [155] Thompson, C. W., Roe, J., Aspinall, P., Mitchell, R., Clow, A., & Miller, D. (2012). More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and urban planning*, 105(3), 221-229.
- [156] Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), 234-240.
- [157] Tseloni, A., & Zarafonitou, C. (2008). Fear of crime and victimization: A multivariate multilevel analysis of competing measurements. *European Journal of Criminology*, 5(4), 387-409.
- [158] Van den Berg, A. E., Maas, J., Verheij, R. A., & Groenewegen, P. P. (2010). Green space as a buffer between stressful life events and health. *Social science & medicine*, 70(8), 1203-1210.
- [159] Wang, H., Calabrese, F., Di Lorenzo, G., & Ratti, C. (2010, September). Transportation mode inference from anonymized and aggregated mobile phone call detail records. In 13th International IEEE Conference on Intelligent Transportation Systems (pp. 318-323). IEEE.
- [160] Weisburd, D., Groff, E. R., & Yang, S. M. (2012). The criminology of place: Street segments and our understanding of the crime problem. Oxford University Press.
- [161] Whitley, R., & Prince, M. (2005). Fear of crime, mobility and mental health in inner-city London, UK. *Social science & medicine*, 61(8), 1678-1688.
- [162] Wikström, P. O. H., & Dolmén, L. (2001). Urbanisation, neighbourhood social integration, informal social control, minor social disorder, victimisation and fear

- of crime. *International Review of Victimology*, 8(2), 121-140.
- [163] Wilcox, P., Quisenberry, N., & Jones, S. (2003). The built environment and community crime risk interpretation. *Journal of Research in crime and delinquency*, 40(3), 322-345.
- [164] Williams, D., Haworth, J., & Cheng, T. (2015, July). Exploratory spatiotemporal data analysis and modelling of public confidence in the police in central London. University of Leeds.
- [165] Williams, D., Haworth, J., & Cheng, T. (2016, April). Predicting public confidence in the police with spatiotemporal Bayesian hierarchical modelling. GIS Research UK (GISRUK).
- [166] Williams, M. L., Burnap, P., & Sloan, L. (2017). Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. *The British Journal of Criminology*, 57(2), 320-340.
- [167] Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic monthly*, 249(3), 29-38.
- [168] Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic monthly*, 249(3), 29-38.
- [169] Wilson, J. S., Kelly, C. M., Schootman, M., Baker, E. A., Banerjee, A., Clennin, M., & Miller, D. K. (2012). Assessing the built environment using omnidirectional imagery. *American journal of preventive medicine*, 42(2), 193-199.
- [170] Woldoff, R. A. (2006). Emphasizing fear of crime in models of neighborhood social disorganization. *Crime Prevention and Community Safety*, 8(4), 228-247.
- [171] Wu, F. (2012). Neighborhood attachment, social participation, and willingness to stay in China's low-income communities. *Urban Affairs Review*, 48(4), 547-570.
- [172] Wyant, B. R. (2008). Multilevel impacts of perceived incivilities and perceptions of crime risk on fear of crime: Isolating endogenous impacts. *Journal of Research in Crime and Delinquency*, 45(1), 39-64.
- [173] Yang, J., Zhao, L., McBride, J., & Gong, P. (2009). Can you see green? Assessing the visibility of urban forests in cities. *Landscape and Urban Planning*, 91(2), 97-104.
- [174] Yao, W., Krzystek, P., & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, 123, 368-380.
- [175] Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., ... & Guan, Q. (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.
- [176] Yuan, Y., & McNeeley, S. (2017). Social ties, collective efficacy, and crime-specific fear in Seattle neighborhoods. *Victims & offenders*, 12(1), 90-112.
- [177] Zeng, L., Lu, J., Li, W., & Li, Y. (2018). A fast approach for large-scale Sky View Factor estimation using street view images. *Building and Environment*, 135, 74-84.
- [178] Zhang, F., Fan, Z., Kang, Y., Hu, Y., & Ratti, C. (2021). "Perception bias": Deciphering a mismatch between urban crime and perception of safety. *Landscape and Urban Planning*, 207, 104003.
- [179] Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, 180, 148-160.

- [180] Zhang, F., Zu, J., Hu, M., Zhu, D., Kang, Y., Gao, S., ... & Huang, Z. (2020). Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems*, 81, 101478.
- [181] Zhang, L., Messner, S. F., Liu, J., & Zhuo, Y. A. (2009). Guanxi and fear of crime in contemporary urban China. *The British Journal of Criminology*, 49(4), 472-490.
- [182] Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 633-641).
- [183] Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2019). Semantic understanding of scenes through the ade20k dataset. *International Journal of Computer Vision*, 127(3), 302-321.
- [184] 崔成, 任红艳, 赵璐, & 庄大方. (2020). 基于街景影像多特征融合的广州市越秀区街道空间品质评估. *地球信息科学学报*, v.22;No.154(06), 158-166.
- [185] 崔喆, 何明怡, & 陆明. (2018). 基于街景图像解译的寒地城市绿视率分析研究——以哈尔滨为例. *中国城市林业*, 16(005), 34-38.
- [186] 杜坤, 宁津生, 田莉, 闫利, & 常坤. (2018). 基于街景影像的交通护栏空间分布特征研究. *地理空间信息*, 016(011), 60-63,71.
- [187] 冯甜甜, & 龚健雅. (2010). 基于建筑物提取的精细尺度人口估算研究. *遥感技术与应用*, 25(003), 323-327.
- [188] 甘欣悦, 余天唯, & 龙瀛. (2018). 街道建成环境中的城市非正规性基于北京老城街景图片的人工打分与机器学习相结合的识别探索. *时代建筑*, (2018年 01), 62-68.
- [189] 韩向娣, 周艺, 王世新, 刘瑞, & 姚尧. (2012). 夜间灯光遥感数据的gdp空间化处理方法. *地球信息科学学报*(01), 128-136.
- [190] 侯杰泰 温忠麟 成子娟. (2006). *结构方程模型及其应用(附光盘)*. 教育科学出版社.
- [191] 胡健波, & 张健. (2018). 无人机遥感在生态学中的应用进展. *生态学报*, 038(001), 20-30.
- [192] 敬峰瑞, 孙虎, & 龙冬平. (2017). 基于网络文本的西溪湿地公园旅游体验要素结构特征分析. *浙江大学学报 (理学版)*, 44(5), 623-630.
- [193] 李宗光, 胡德勇, 李吉贺, & 岑建. (2016). 基于夜间灯光数据的连片特困区gdp估算及其空间化. *国土资源遥感*, 28(002), 168-174.
- [194] 廖顺宝, & 孙九林. (2010). 基于 GIS 的青藏高原人口统计数据空间化. *地理学报*, 58(1), 25-33.
- [195] 廖薇薇, 柳林, 周素红, 宋广文, 李秋萍, & 刘凯. (2018). 多源在线地理编码服务分类优化模型. *热带地理*.
- [196] 林爱华、乔红芳、沈利生. (2020). 环境规制对绿色发展水平的作用效果——基于断点存在时的面板被调节中介分析. *资源开发与市场*, v.36;No.280(12), 25-33+100.
- [197] 林冬阳, & 田宝江. (2014). 犯罪被害恐惧与城市空间环境关系的研究综述. *城乡治理与规划改革——2014中国城市规划年会*.
- [198] 刘云舒, 赵鹏军, & 梁进社. (2018). 基于位置服务数据的城市活力研究——以北京市六环内区域为例. *地域研究与开发*, 6.
- [199] 罗皓文. (2019). 大尺度城市区域情感感知度量及空间异质性研究. (Bachelor's dissertation, 中山大学)

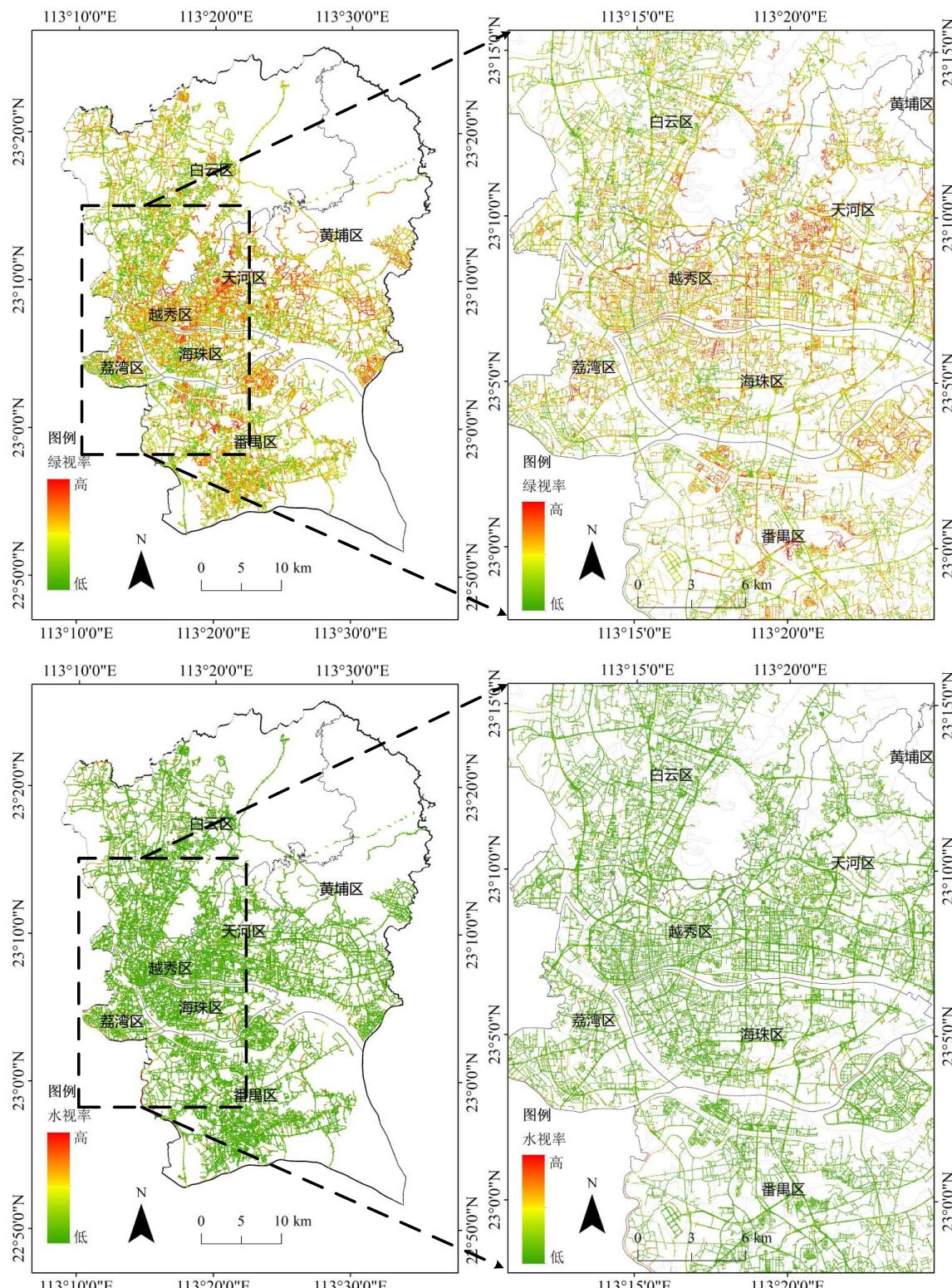
- [200] 钮心毅, 王垚, & 丁亮. (2017). 利用手机信令数据测度城镇体系的等级结构. 规划师, 33(1), 50-56.
- [201] 彭玉伟. (2015). 老年人高犯罪恐惧与低被害风险的矛盾. 中国老年学杂志, 24.
- [202] 青木陽二. (1987). 視野の広がりと緑量感の関連. 造園雑誌, 51(1), 1-10.
- [203] 唐婧娴, & 龙瀛. (2017). 特大城市中心区街道空间品质的测度——以北京二三环和上海内环为例. 规划师, 033(002), 68-73.
- [204] 王德, 钟炜菁, 谢栋灿, & 叶晖. (2015). 手机信令数据在城市建成环境评价中的应用——以上海市宝山区为例. 城市规划学刊, (2015年05), 82-90.
- [205] 王德, 朱查松, & 谢栋灿. (2016). 上海市居民就业地迁移研究——基于手机信令数据的分析. 中国人口科学, 000(001), 80-89.
- [206] 王珂. 城市公众安全感的影响因素研究. (Doctoral dissertation, 华南理工大学).
- [207] 王娅君. 东南亚若干首都城市街景图像汉字检测与时空分布分析. (Doctoral dissertation, 南京大学).
- [208] 王昱. (2016). 基于车载LiDAR数据和街景照片的街道美景度评价. (Doctoral dissertation, 南京大学).
- [209] 韦雪花, 王永国, 郑君, 王萌, & 冯仲科. (2013). 基于三维激光扫描点云的树冠体积计算方法. 农业机械学报, 44(7), 235-240.
- [210] 魏书法, & 程章林. (2017). 基于图像的城市场景垃圾自动检测. 集成技术, (2017年01), 39-52.
- [211] 温忠麟, & 叶宝娟. (2014). 中介效应分析: 方法和模型发展. 心理科学进展, 22(005), 731-745.
- [212] 温忠麟, 刘红云, & 侯杰泰. (2012). 调节效应和中介效应分析:#Analyses of moderating and mediating effects. 教育科学出版社.
- [213] 温忠麟, 张雷, 侯杰泰, & 刘红云. (2004). 中介效应检验程序及其应用. 心理学报, 36(05), 614-620.
- [214] 吴健生, 许多, 谢舞丹, & 彭建. (2015). 基于遥感影像的中尺度人口统计数据空间化——以京津冀地区为例. 北京大学学报(自然科学版)(04), 707-717.
- [215] 肖鸿元. 基于地点的城市社区居民安全感影响因素及作用机制分析. (Doctoral dissertation, 西安外国语大学).
- [216] 姚海芳, 冯天楠, & 刘劲松. (2020). 基于手机信令数据的机场航空旅客分类识别研究——以石家庄正定国际机场为例. 地理与地理信息科学, v.36(03), 62-68.
- [217] 叶雷. (2017). 机器学习算法在医疗数据分析中的应用. (Doctoral dissertation, 华中师范大学).
- [218] 叶宇, 张昭希, 张啸虎, & 曾伟. (2019). 人本尺度的街道空间品质测度——结合街景数据和新分析技术的大规模、高精度评价框架. 国际城市规划, 34(01), 18-27.
- [219] 张鹏程, 赵齐, & 高泽宇. (2019). 一种融合移动边缘计算和深度学习的城市街道垃圾检测和清洁度评估方法. 小型微型计算机系统, 4.
- [220] 张延吉, 秦波, & 唐杰. (2017). 城市建成环境对居住安全感的影响——基于全国278个城市社区的实证分析. 地理科学(09), 1318-1325.
- [221] 张延吉, 秦波, & 朱春武. (2019). 北京城市建成环境对犯罪行为和居住

- 安全感的影响. 地理学报, 74(2), 238-252.
- [222] 张永霖,付晓. (2020). 基于深度学习街景影像解译和景感生态学的视域环境定量解读. 生态学报, v.40(22), 183-190.
- [223] 赵鹏军, & 曹毓书. (2018). 基于多源lbs数据的职住平衡对比研究 ——以北京城区为例. 北京大学学报(自然科学版), 54(06), 157-169.
- [224] 钟炜菁, 王德, 谢栋灿, & 晏龙旭. (2017). 上海市人口分布与空间活动的动态特征研究: 基于手机信令数据的探索. 地理研究, 36(5), 972-984.
- [225] 朱志玲. (2020). 警察素质,警察行为与治安安全感——基于cgss2015数据的实证分析. 贵州警察学院学报(4).

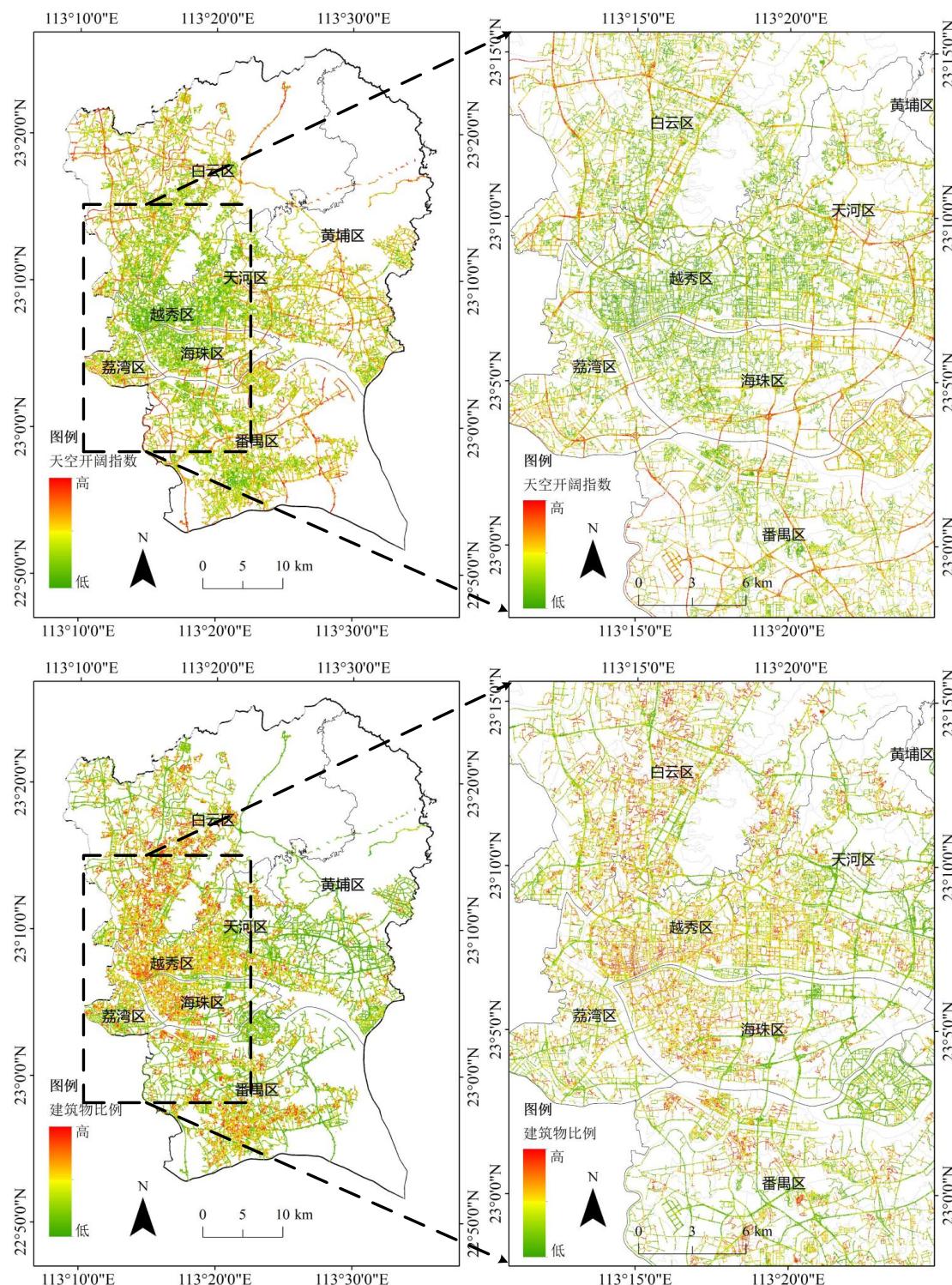


## 附录

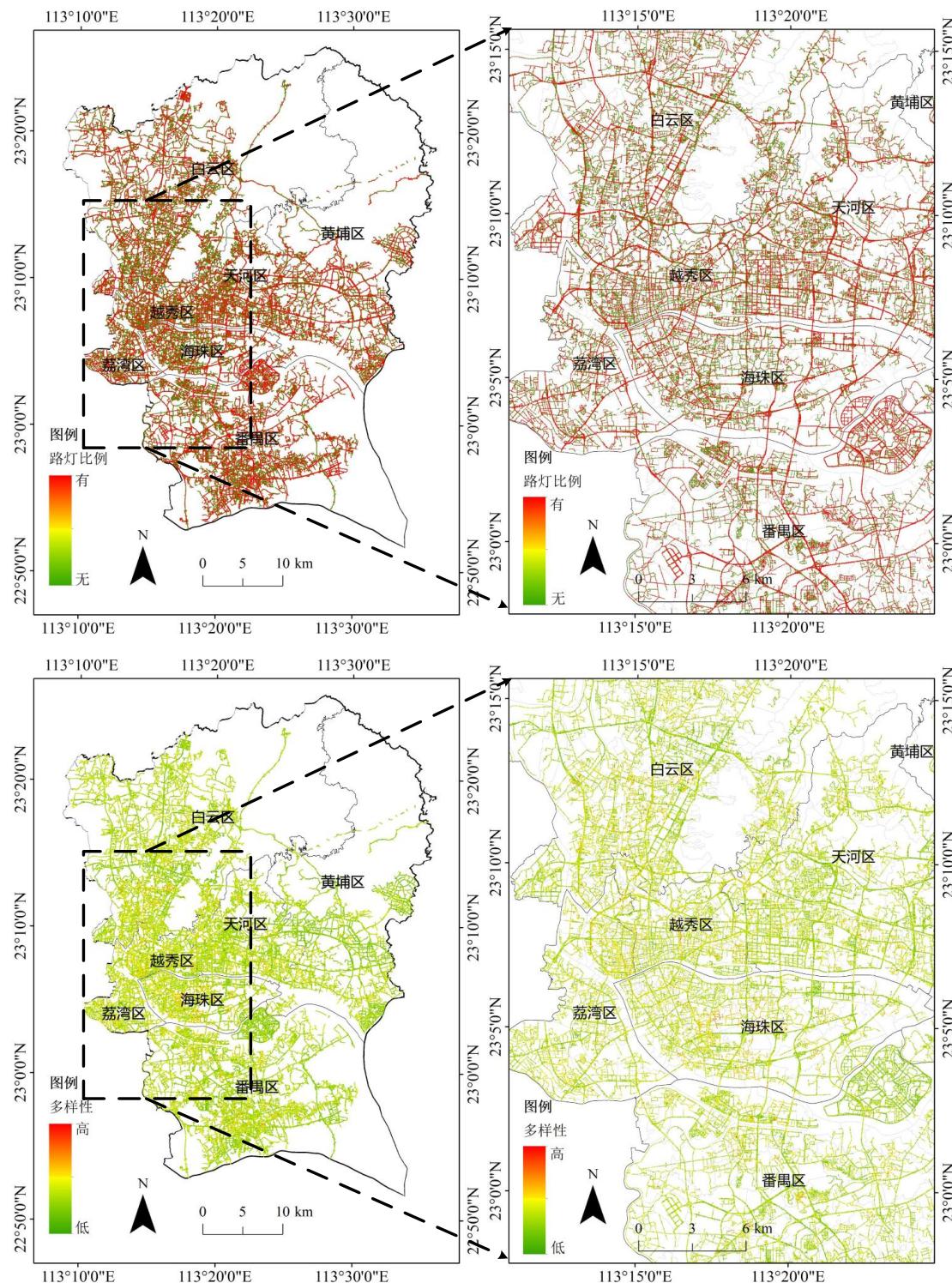
附表 1 点状尺度的客观街景微环境空间分布



附表图 A 绿视率和水视率微环境要素空间分布结果



附表图 B 天空开阔程度和建筑物比例微环境要素空间分布结果



附表图 C 路灯比例和多样性微环境要素空间分布结果

附表 2 基于调研社区的预测值和真实值不安全感对照表

序号	社区	真实值	预测值	绝对误差
1	罗溪村委会	0.25	0.715897	0.465897
2	草庄村委会	0.24	0.434406	0.194406
3	洛溪新城社区居委会	0.142857	0.101357	0.0415
4	东兴社区居委会	0.1	0	0.1
5	北亭村民委员会	0.16	0.013126	0.146874
6	厦滘村民委员会	0.15	0.012308	0.137692
7	沙坛一村民委员会	0.1	0.03329	0.06671
8	小鸟村民委员会	0.35	0.26956	0.08044
9	社学社区居委会	0	0	0
10	愉景雅苑社区居委会	0.05	0.053008	0.003008
11	怡乐社区居委会	0.166667	0.406446	0.239779
12	晓港东社区居委会	0.15	0.544885	0.394885
13	堑口社区居委会	0.095238	0.103167	0.007929
14	江宝社区居委会	0.45	0.527887	0.077887
15	敬和里社区居委会	0.1	0	0.1
16	嘉仕社区居委会	0.083333	0.021017	0.062316
17	海南社区居委会	0.05	0	0.05
18	西关大屋社区居委会	0.2	0.158958	0.041042
19	小策社区居委会	0.25	0.660071	0.410071
20	河傍社区居委会	0.095238	0	0.095238
21	杉栏社区居委会	0.5	0.934467	0.434467
22	长平社区居委会	0.1	0	0.1
23	金坑村委会	0.363636	0.613024	0.249388
24	东晖社区居委会	0.5	0.987189	0.487189
25	绿洲社区居委会	0.24	0.174333	0.065667
26	暨大社区居委会	0.12	0.069034	0.050966
27	华农社区居委会	0.08	0.000659	0.079341
28	恒怡社区居委会	0.05	0.024022	0.025978
29	水荫四横路社区居委会	0.25	0.110708	0.139292
30	文昌社区居委会	0.1	0.082358	0.017642

## 附录

31	猎德社区居委会	0.3	0.565767	0.265767
32	利民社区居委会	0.05	0.012398	0.037602
33	天誉社区居委会	0.095238	0.323102	0.227864
34	马棚岗社区居委会	0.142857	0.005382	0.137475
35	海印社区居委会	0.25	0.574816	0.324816
36	果菜西社区居委会	0.25	0.359364	0.109364
37	石将军社区居委会	0.55	0.551308	0.001308
38	和义巷社区居委会	0.3	0.492035	0.192035
39	二马路社区居委会	0.1	0.089336	0.010664
40	明月社区居委会	0.052632	0	0.052632
41	张村社区居委会	0.1	0.024462	0.075538
42	螺涌社区居委会	0.4	0.49896	0.09896
43	平沙南社区居委会	0.25	0.304189	0.054189
44	江夏社区居委会	0.190476	0.315589	0.125113
45	诜村村民委员会	0.2	0.393538	0.193538
46	龙美村民委员会	0.190476	0	0.190476
47	上教村民委员会	0.238095	0.149501	0.088594
48	上涌东约社区居委会	0.541667	0.542113	0.000447
49	康乐西社区居委会	0.411765	0.244663	0.167102
50	姬堂社区居委会	0.25	0.284449	0.034449
51	下沙社区居委会	0.238095	0.208353	0.029742
52	夏园社区居委会	0.142857	0.321564	0.178707
53	茅岗社区居委会	0.238095	0.534899	0.296803
54	东望社区居委会	0.380952	0.249132	0.131821
55	河沙社区居委会	0.5	0.793127	0.293127
56	笔岗社区居委会	0.25	0.663632	0.413632
57	怡东社区居委会	0.05	0	0.05
58	兴华社区居委会	0.095238	0.172474	0.077236
59	程界西社区居委会	0.210526	0.249672	0.039146
60	王圣堂社区居委会	0.25	0.10097	0.14903
61	聚德东社区居委会	0.275862	0.215954	0.059908
62	平乐社区居委会	0.285714	0.113984	0.17173
63	龙归社区居委会	0.241379	0.356305	0.114926
64	泽德社区居委会	0.133333	0.161273	0.027939

## 附录

65	棠德北社区居委会	0.115385	0.006252	0.109133
66	东漖社区居委会	0.44	0.498751	0.058751
67	西滘社区居委会	0.217391	0.472133	0.254742
68	吉山西社区居委会	0.291667	0.530104	0.238438
69	仰忠社区居委会	0.285714	0.61976	0.334046
70	雅荷塘社区居委会	0.25	0.439875	0.189875
71	三元巷社区居委会	0.3	0.348426	0.048426
72	文园巷社区居委会	0.35	0.310174	0.039826
73	耀华社区居委会	0.65	0.941997	0.291997
74	冼基社区居委会	0.15	0.027849	0.122151
75	新风社区居委会	0.05	0	0.05
76	龙津东社区居委会	0.3	0.38061	0.08061
77	中市社区居委会	0.65	1	0.35
78	永龙社区居委会	0.25	0.48771	0.23771
79	华苑社区居委会	0.590909	0.799083	0.208174
80	得胜社区居委会	0.65	0.992092	0.342092

## 博士期间发表的学术论文、参与的科研项目和获得的奖励

已发表相关学术论文：

1. **Jing, F.**, Liu, L., Zhou, S., Song, J., Wang, L., Zhou, H., ... & Ma, R. (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. (与本文第四章、第五章、第六章相关，影响因子 2.849、SSCI JCR 二区、第一作)。
2. **Jing, F.**, Liu, L., Zhou, S., & Song, G. (2020). Examining the Relationship between Hukou Status, Perceived Neighborhood Conditions, and Fear of Crime in Guangzhou, China. Sustainability, 12(22), 9614. (与本文第二章、第五章相关，影响因子 2.576、SCI /JCR 二区、第一作)。
3. Zhou, H; Liu, L; Lan, M; Zhu, W; Song, G; **Jing, F**; Zhong, Y; Su, Z; Gu, X. (2021). Using Google Street View imagery to capture micro built environment characteristics in drug places, compared with street robbery. Computers, Environment and Urban Systems, 88, 101631. (与本文第三章相关，影响因子 4.655、SCI /JCR 一区、第六作)。
4. Long, D., Liu, L., Feng, J., Zhou, S., & **Jing, F.** (2018). Assessing the influence of prior on subsequent street robbery location choices: a case study in ZG city, China. Sustainability, 10(6), 1818. (影响因子 2.576、SSCI JCR 二区、第五作)。

待发表相关学术论文：

1. **Jing, F.**, Liu, L., & Zhou, S. (2021). How does Crime-specific Victimization Impact Fear of Crime in China? The Role of Neighborhood Characteristics. International Journal of Offender Therapy and Comparative Criminology. Recommended publication after revisions.

学术交流：

2019.09-2020.03 赴英国 Manchester Metropolitan University 犯罪与幸福感大数据中心访学

相关科研项目：

1. 参与国家重点研发项目“城镇公共安全立体化网络构建与应急响应示范”的子

项目“大数据安全感与应急响应”(2018-07 to 2021-06) (Nos. 2018YFB0505500, 2018YFB0505503)

2. 参与由广州市政法委资助的“社会安全评估体系研究项目”(2018-09 to 2018-12)

**专利:**

柳林, 敬峰瑞, 周素红, 刘凯等(2019). 一种街道环境多维度失序的空间分布识别方法. 公开号: CN111008657A 审中

## 致谢

广州，再见。求学，是不断告别的历程。小学四年级结束，满怀憧憬地告别了真正意义上的故乡—蒙垭庙村。小学六年级结束，告别了升水镇。高中结束，告别了南部县。大学结束，告别了武汉。硕士结束，告别了西安。博士结束，继续告别，再见广州。人生的前三分之一，不断在告别，不断在留恋。远方，总是那么令人憧憬，又总是那么令人迷茫。

很幸运，兴趣成为了引导我前进的航标。来自那么小的地方，自然好奇那么大的世界。地图，给了我认识世界的窗口。记得小时候，看着中国地图下面的“泰国”，我认成“秦国”，一直奇怪怎么古代的秦国还存在呢？这成了我小时候关于地图的最深刻印象。中学时，我会在新华书店，翻阅国外城市地图册，看着巴黎、柏林这些知名而又遥远的城市，让人着迷。高中时，枯燥的学习之余，看世界地图就成了很大的快乐。地理，不知不觉，成为了我的重要兴趣。大学时，因为文科生不能选择地理专业，后来学了社会工作和心理学。大三，当面临考研时，发现自己还是对地理学格外有兴趣，就做出了跨考的决定。兜兜转转，到陕师大学习自然地理学，也正式开始了自己的城市研究之路。读博时，选择了地图学与地理信息系统专业。就这样，和地理的交集越来越深。和大多数人从理工到文科、商科相反，我从文科逐渐跨到理工科。这一路，的确很多不容易。不过，地理就是结合自然科学和社会科学的一个学科，所以我的背景好像也还适合。

博士学位论文的完成，要感谢很多人。感谢柳林老师。柳老师对待学生平易近人，始终支持、相信、鼓励学生，认真指导学生科研，想方设法让学生顺利毕业。正是在这样的环境下，我一直能够以兴趣驱动科研，还愿意接下来继续从事科研。特别地，几件事情上想感谢柳老师。第一，感谢柳老师在博士生面试的时候选择了我。记得面试时还有一位本硕都是中大的学生也报考，所以我也很幸运。第二，感谢柳老师帮忙联系外导，同意我去英国访学，以及给予机会参加美国 AAG 年会，让我见识到了外面更大的世界。第三，在我去年春天想放弃的时候，柳老师没有责骂，而是理解，给予我缓冲的空间。所以，非常感谢柳老师。

感谢周素红老师。记得最开始联系柳老师的时候，柳老师让我发邮件由周老师代为审核。读博期间，在项目上、科研上和周老师都保持着密切联系，周老师也指导了我的整个博士阶段。感谢周老师，总是耐心地和我讨论我遇到的各种科研问题，助我打开思维，带我参加组会，让我受益匪浅。即使我在英国访学时，也会关心我的各种生活适应问题。在去年我想要放弃的时候，也是一直支持、鼓励、帮助我，最终助我走出困境。

感谢外导 Jon Bannister 教授。访学前，Jon 和我一起完成研究计划，帮助我联系研

究生院，办理各种材料。访学时，也定期和我交流，给予我宽松环境，让我在 BDC 大家庭很快适应。回国后，当我遇到各种疑惑的时候，也都耐心帮助我。

感谢参加我预答辩和答辩的老师们，包括黎夏老师、张新长老师、刘小平老师、石茜老师、陈楚群老师、王兆礼老师、李军老师。感谢在我填写访学材料、推荐材料等方面提供了不少帮助的综合地理信息中心的老师们，包括刘凯老师、王先伟老师、李秋萍老师、贺智老师、梁育填老师、习倩晴老师等。感谢中国地质大学（武汉）的姚尧老师，在模型运用和文章推荐上对我很大帮助。特别感谢广州大学的宋广文老师和肖露子老师，我整个博士阶段都给予我热心、真诚的指导，好像我在广州的家人。同样，还要特别感谢请我在中大吃第一顿饭的龙冬平老师、和我很有共同观念的陈建国老师。

感谢帮助过我的朋友们。感谢师门的冯嘉欣（经常一起探讨学术问题）、余洪杰（一个很好的让我心态平和的参照对象）、廖薇薇（给予我很多技术指导）、张政（会玩、有趣的师妹）、梁斯毅（热情、真诚的师弟）、孙秋远（广州大学联络员）等同学。感谢周老师师门的周淑丽（经常一起约饭和生活交流）、卢俊文（从探讨街景问题到手机数据运用上的很多帮助）、张雪（一起沟通了很多博士生活问题）、陈菲（学术、生活上都有帮助）、郑重老师（组会上的意见帮助多多）等。感谢胡洁（生活建议上给予了诸多帮助）、黎丰收（在鹭江站接我加入团队，雅思好帮手）、王逸文（重要的技术问题沟通对象）、许贤鹏、刘文雅、以及其他 17 级的硕士同学等。感谢一些博士同窗同学，包括学习生活中的好友方勇军，答辩路上帮助很多的张大川，深夜开会好搭档阮永俭和张绍云、期盼我回成都发展的老乡石洪、老友袁超、聚餐好友张子昂、严林、张瑞豪等同学。感谢在英国访学期间的团队同事Kitty, WonDo, Monsuru, Karolina, Samuel, Salman, Mark 等博士，以及程建权和段炼老师、黄蓉、尹超同学，好友 Dario、胡开，以及 Joel 等教会朋友们（给予了我锻炼英语、适应文化的平台）。感谢目前在北美的周翰林（探讨了很多学术问题）、顾欣同学。特别地，要感谢在方法上指导了我的同学们，包括经常请教技术问题的梁燊、深度学习方面有很大帮助的张森源、罗皓文等。也特别感谢论文数据处理上帮助了很多的同学们，包括随叫随到的林荣平博士，事业有成的宋江宇、执行能力强的王林森、技术能手马若飞。感谢的人还有很多，就不再一一叙述。

感谢一直以来给我帮助的亲友们（如杜桂清、杜德映、敬丹霞、敬智雄、闵宝财、敬守宏、杜维雄、敬守东、敬守英、敬仕俊等）。我们大家庭里，好像没有谁读到了博士。大家眼中，我成了很聪明的人。最后，特别感谢一些特别的人。感谢我的师妹张卓芳，求学最后的一年里，遇到了可以一直走下去的人。感谢我的家人，敬守林、杜玉华、敬静，不论什么时候，你们都给我无条件的支持。

也许很多年后的某一天，我会回想起来，完成博士论文的这一天，很普通。

