

# 도시 기능구역별 운전자의 가시적 경관이 자동차 교통사고에 미치는 영향 분석\*

: 해석가능한 기계학습과 음이항 회귀모형의 혼합적 접근을 중심으로

## Analysis of the Impact of Driver's Visual Landscape on Traffic Accidents in Urban Functional Zones

: A Hybrid Approach Combining Interpretable Machine Learning and Negative Binomial Regression

문정훈\*\* · 이수기\*\*\*

Moon, Jeonghun · Lee, Sugie

### Abstract

Urban functional zones exhibit distinct patterns of traffic accidents, reflecting variations in spatial configuration and visual landscapes. This study investigates how drivers' visual landscape elements influence automobile traffic accidents in Seoul. Based on 551,804 Point of Interest (POI) records obtained from Kakao Maps, the city was categorized into four functional zones using Word2Vec embeddings and K-means clustering: high-density residential and educational areas; urban mixed-use activity areas; transit and leisure mixed-use areas; and low-density residential and local living areas. Visual landscape indices, including the Green View Index (GVI), Color Entropy Index (CEI), and Visual Obstruction Index (VOI), were derived from 95,842 Naver Street View images through semantic segmentation and color analysis. XGBoost with SHAP values was used to identify influential variables, followed by negative binomial regression for statistical validation. The results indicate that CEI reduces accidents in high-density residential and educational areas; GVI reduces accidents in urban mixed-use activity areas; VOI increases accidents in transit and leisure mixed-use areas; and GVI increases accidents in low-density residential and local living areas. These findings suggest that the effects of visual landscape elements differ across functional zones, underscoring the need for tailored urban design and traffic safety strategies.

**주제어** 도시 기능구역, 교통사고, Word2vec, 해석 가능한 기계학습, 음이항 회귀모형

**Keywords** Urban Functional Zones, Traffic Accidents, Word2vec, Explainable Machine Learning, Negative Binomial Regression

## 1. 연구의 배경 및 목적

도시는 상업·주거·산업 등 다양한 기능구역으로 분화되어 있으며, 각 구역은 고유한 물리적·사회적·기능적 특성을 내포하고 있다. 이러한 공간적 이질성(Spatial Heterogeneity)은 동일한 도시 내에서도 교통사고의 발생 빈도와 위험 요인이 상이하게 나

타나는 주요 원인으로 지적된다(Sun et al., 2018). 예를 들어, 상업 지역의 경우 차량과 보행자의 상호작용이 빈번하게 발생하는 반면, 주거 지역은 협소한 도로 폭과 무질서한 주차 문제 등 다른 유형의 안전 문제가 나타날 수 있다. 그럼에도 불구하고, 기존의 교통사고 연구는 도시 전체를 균질한 공간으로 가정하여 주로 운전자 개별 특성 및 도로의 물리적 구조에 초점을 맞추는 경향이

\* 본 연구는 2024년 대한국토·도시계획학회 추계학술대회에서 발표한 우수논문상 수상 논문을 수정·보완하여 작성하였음.

\*\* Master Student, Department of Urban Planning & Engineering, Hanyang University (First Author: littledoor@hanyang.ac.kr)

\*\*\* Professor, Department of Urban Planning & Engineering, Hanyang University (Corresponding Author: sugielee@hanyang.ac.kr)

강해, 이러한 기능구역별 환경 특성을 충분히 반영하지 못하였다. 따라서 각 구역의 고유 특성을 고려한 차별화된 분석과 정적 접근이 필요하다.

한편, 최근 연구에서는 광고판, 간판, 주변 건축물과 같은 가시적 경관 요소가 운전자의 주의력 분산과 인지 부하를 통한 정보 처리 지연을 유발함으로써 교통사고 발생 위험을 높일 수 있음을 보고하고 있다(Regan et al., 2011; Dukic et al., 2013; Chiang et al., 2022). 이러한 점에서 가시적 경관 요소는 전통적으로 논의되어 온 운전자의 인적 특성이나 도로의 물리적 구조와 더불어, 교통사고 발생을 설명하는 데 중요한 고려 요인으로 부상하고 있다. 그러나 동일한 시각적 요소에 대해서는 연구마다 상이한 결과가 보고되고 있으며, 이러한 불일치의 구조적 원인에 대한 체계적 분석은 부족한 실정이다.

이에 본 연구는 도시 기능구역의 밀도, 용도, 배치 등 공간적 특성과 가시적 경관 요소 간 상호작용이 교통사고 발생에 미치는 영향 차이를 야기할 수 있다는 가능성에 주목하여 다음 두 가지 핵심 연구질문을 설정하였다. 첫째, “실제 도시 활동 특성을 반영한 기능구역은 어떠한 유형으로 구분되는가?” 둘째, “유형화된 도시 기능구역별 가시적 경관 요소가 교통사고 발생 빈도에 어떠한 차별적 영향을 미치는가?”이다. 이러한 질문에 답하기 위해, 본 연구는 서울시를 대상으로 기능구역을 실증적으로 유형화하고, 각 구역별 가시적 경관 요소가 자동차 교통사고 발생에 미치는 차별적 효과를 계량적으로 검증하는 것을 목적으로 한다.

## II. 선행연구 검토

### 1. 도시 기능구역 분류 연구

먼저, 도시 기능구역(Urban Functional Zones, UFZs)이란 도시 내부에서 주거, 상업, 산업, 공공서비스 등 다양한 기능을 수행하는 공간적 단위를 의미한다(Lie et al., 2023; 2025). UFZ 개념은 도시공간의 구조적·기능적 특성을 체계적으로 파악하고, 이에 기초한 도시계획 및 관리를 도모하기 위한 핵심 연구 기반으로 간주된다.

초기 연구는 주로 생태학 및 지리학적 관점에서 출발하여, 자연 생태계 보존과 생물 다양성 유지를 위한 공간구조의 분화에 주목하고(Liu et al., 2021) 주거, 상업, 산업, 공공서비스, 녹지 및 교통 등 법적 토지이용에 근거한 범주화가 주류를 이루었다(Chi et al., 2016; Kang et al., 2018; Liu et al., 2021). 이러한 분류는 거시적 기능 분포를 파악하는 데는 유용하지만, 최근 급격한 도시화로 인한 공간적 배치와 기능적 상호작용이 다양해지면서, 도시의 실질적 활동 및 역동성을 세밀하게 포착하는 데 한계가 있다.

이러한 한계를 극복하기 위해 최근에는 관심지점(Point of Interest, POI) 데이터를 활용하는 데이터 기반 접근이 확산되고

있다(Zhai et al., 2019). POI는 인간 활동의 실제 중심지와 공간적 패턴을 반영하므로, 도시 기능구역 도출을 위한 핵심적인 정량적 지표로 주목받고 있다.

나아가, 최근에는 POI의 공간적 분포와 기능적 의미를 함께 반영하기 위해 기계학습 기반 임베딩(Word2vec, Doc2vec)과 토픽 모델링(Latent Dirichlet Allocation, LDA) 등이 도입되고 있다(Lin et al., 2024). 특히, Word2vec은 도시 내 활동 간의 공간적·의미적 관계를 고차원 벡터 공간에서 구조화하는 데 강점을 보인다. 국내에서도 서울시를 사례로 중심지 도출과 기능 유형화를 시도하거나(김선재·이수기, 2021), 세종시에서 지역별 기능 특성을 규명하는 등(이영민, 2022) 데이터 기반 방법론의 적용 가능성과 타당성이 확인되고 있다. 또한 임베딩 기반 유형화 시도도 이루어지며(이지원·김지영, 2023), 국내 연구 적용의 실효성을 뒷받침한다. 따라서 본 연구는 POI 데이터와 Word2vec 임베딩 기법을 활용하여 서울시 내 다양한 기능과 활동의 실질적 관계를 반영하는 도시 기능구역을 정의하고자 한다.

### 2. 자동차 교통사고와 가시적 경관 요소의 연관성 연구

자동차 교통사고는 도로 환경의 물리적 특성과 운전자 요인의 상호작용으로 발생한다. 인지 요인이 주된 설명력으로 알려져 있으나, 주행 중 노출되는 가시적 경관 요소 또한 유의한 환경 변수로 논의되고 있다(Cai et al., 2022). 이는 운전자의 PIEV 인지반응 과정(Perception-Identification-Emotion-Volition) 중 초기 단계인 지각과 판별 과정에 개입하여 주의를 분산시키고 정보 처리 속도를 저하시킴으로써 반응 시간을 지연시키는 직접 경로로 사고 위험에 작동할 수 있기 때문이다. 이러한 메커니즘은 도로 주변의 시야 개방성(혹은 폐쇄성), 색상 복잡도, 도로변 녹지 등 시각 환경 지표와 사고 발생 간 유의한 연관성이 확인되었다(Choi et al., 2020). 다만 그 효과의 크기와 방향성은 ‘일률적’이 기보다, 맥락에 따라 상반된 결과로 보고되었다.

시야 폐쇄성의 경우, 공간적 경계를 명료하게 인식시켜 주의를 환기를 유도하는 보호 효과가 보고된 반면(Harvey and Aultman, 2015), 원거리 가시성을 제한하여 사고 위험을 증가시키는 결과도 제시된다(Kwon and Cho, 2020; Stiles et al., 2022). 또한, 색상 복잡도의 경우 이중적 양상이 관찰되었다. 다양한 색채와 높은 대비가 시각적 경계심과 환경 인지를 강화하여 사고 위험을 완화한다는 근거가 있는 한편(박준상, 2023), 과도한 시각적 피로나 인지 과부하를 야기하여 사고 위험을 높이는 경향이 확인된다(Zhang et al., 2023). 마지막으로 도로변 녹지의 경우, 가로수와 수목이 운전자의 시각적 피로 완화와 과속 경향을 억제하여 사고 감소와 연관된다는 결과가 다수 보고되었으나(Nitoslawski et al., 2016; Song et al., 2018; Van Treese et al., 2018), 일각에서는 녹지의 밀집 및 배치가 시야를 차단하거나 주의를 분산시켜

사고 위험을 증가시킬 수 있다는 상반된 결과도 제시된다(Cai et al., 2022; Choi et al., 2020; Han, 2025).

이처럼 동일한 가시적 경관요소 임에도 불구하고, 효과의 방향성과 효과 크기가 맥락에 따라 달리 추정되는 현상이 반복적으로 관찰된다. 그러나 이러한 불일치의 구조적·맥락적 원인에 대한 체계적 분석은 미흡한 실정이다. 따라서 본 연구는 기존 선행연구에서 주목한 폐쇄성(Enclosure), 색상 복잡도(Color Complexity), 녹지 지수(Green View Index)를 핵심 변수로 설정하여, 기능구역별 차별적 영향을 실증적으로 규명하고자 한다.

### 3. 교통사고 분석 연구 방법론

교통사고 발생 요인을 규명하는 방법론 측면에서, 이산형 빈도 자료 특성에 부합하는 전통적 통계모형이 주로 사용되어 왔다. 특히 사고 건수 자료에는 포아송 회귀모형(Poisson Regression)이 기본적으로 적용되며(Ye et al., 2018), 과대산포(Overdispersion)가 존재할 경우, 이를 효과적으로 제어할 수 있는 음이항 회귀모형(Negative Binomial Regression)이 널리 사용된다(Dong et al., 2015; Yang et al., 2021). 이러한 회귀모형은 변수와 사고 발생 간 통계적 유의성을 검증하는 데 유용하지만, 계수의 크기가 측정 단위, 분산, 범위에 좌우되고, 표준화 계수(Standardized Coefficient)를 적용하더라도 다중공선성, 변수 간 상호작용, 비선형성 등으로 인해 변수 간 상대적 기여를 일관되게 비교하기에는 제약이 존재한다.

최근 이러한 한계를 극복하고자 기계학습 기반 분석기법이 교통사고 연구에 보완적으로 도입되고 있다. 특히, 기계학습 알고리즘 중 eXtreme Gradient Boosting(XGBoost)는 높은 예측 정확도와 복잡한 변수 간 비선형 관계를 잘 포착할 수 있다는 점에서 교통사고 발생 분석에서 강점을 나타내고 있다(Chang et al., 2022; Nassiri et al., 2023). 더불어 SHapley Additive exPlanations(SHAP) 기법과의 결합을 통해, 복잡한 모형 내 변수 기여도를 정량적으로 평가하고, 그 결과를 시각적으로 제시함으로써 직관적으로 해석할 수 있는 분석 기반을 마련하고 있다(Parsa et al., 2020; Aboulola, 2024).

그러나 XGBoost-SHAP 분석 역시, 변수의 통계적 유의성 검증이나 인과관계 해석 측면에서는 한계가 존재한다. 따라서 본 연구는 XGBoost-SHAP 기법을 통해 상대적으로 중요도가 높은 핵심 변수를 선별하고, 이를 바탕으로 전통적 통계모형을 적용하여 각 변수의 효과 부호와 통계적 유의성을 교차 검증하였다. 이러한 혼합적 접근은 예측 성능에 근거한 변수 중요도를 활용해 가시적 경관 요소의 정량적 위상을 기존 인적·물리적 요인과 비교하고, 통계적 검증으로 분석 결과의 신뢰성을 확보하고자 한다.

## 4. 연구의 차별성

본 연구는 기존 선행연구의 한계를 보완하고, 도시공간 구조와 교통사고 발생 메커니즘 간의 연계를 규명하기 위해 다음 세 가지 측면의 학술적 차별성을 지닌다.

첫째, 기존의 행정구역 단위 또는 정형화된 토지이용 기반의 기능구역 분류 방식과 달리, POI 및 임베딩 기법(Word2vec)을 적용함으로써, 실제 도시 내 활동의 공간적 패턴과 의미론적 유사성을 반영한 기능구역을 도출하였다. 이를 기반으로 교통사고 발생 특성과의 연계 분석을 수행함으로써, 도시 공간의 기능성과 안전성 간의 관계를 구역별로 세분화하여 접근한 점에서 기존 연구와 구분된다. 둘째, 시각환경 요소를 정량화하는 데 있어, 본 연구는 Naver Street View 이미지 데이터를 활용하여 운전자 시점에서 인지 가능한 경관 요소(시야 차단성, 색상 복잡도, 도로 변 녹지 등)를 추출하였다. 이는 기존 연구들이 2D 공간자료에 의존했던 분석 한계를 넘어, 실제 운전환경에 기반한 시각적 정보에 초점을 맞추었다는 점에서 분석 관점의 차이를 가진다. 셋째, 분석기법에 있어 기계학습 모형과 전통적 통계모형을 병행 적용함으로써, 변수의 상대적 영향력과 통계적 유의성을 동시에 평가하였다. 이러한 방법론적 구성은 기존 단일 접근 방식과 차별화된다.

## III. 분석의 틀

### 1. 사례 지역 및 연구의 범위

본 연구의 공간적 범위는 서울특별시 전역이다. 서울특별시는 우리나라의 대표적인 대도시로, 전국에서 가장 많은 인구와 차량 등록 대수를 보유하고 있어 교통사고 발생 빈도 또한 매우 높다(서지민·이수기, 2016). 또한, 서울시는 다핵·분산형 공간구조를 갖추고 있어 다양한 활동 중심지가 분포하고 있는 지역이다(김선재·이수기, 2021). 이러한 특성은 도시의 공간구조와 교통사고 패턴 간의 관계를 분석하기에 적합한 사례 지역이라 할 수 있다. 분석 단위로는 500m×500m 격자를 선정하였다. 격자 단위는 공간을 균등하게 분할할 수 있어 분석의 일관성을 확보하고, 결과의 해석이 용이한 장점이 있다(Liu et al., 2021). 그중 500m 격자는 이지원·김지영(2023)의 연구에서 도시 기능구역을 정의하는 데 적절한 단위로 검증되었으며, 교통사고 분석에서도 도로 환경의 동질성을 평가하는 데 유효한 분석 단위로 확인되었다. 연구의 시간적 범위는 2017년부터 2019년까지로 설정하였다. 이는 코로나19 팬데믹의 영향으로 인한 교통 패턴의 변화를 최소화하기 위한 조치로, 팬데믹 이전의 일반적인 교통사고 패턴을 반영할 수 있는 적절한 기간으로 판단되었다. 사례 지역 및 공간적 범위는 <그림 1>과 같다.

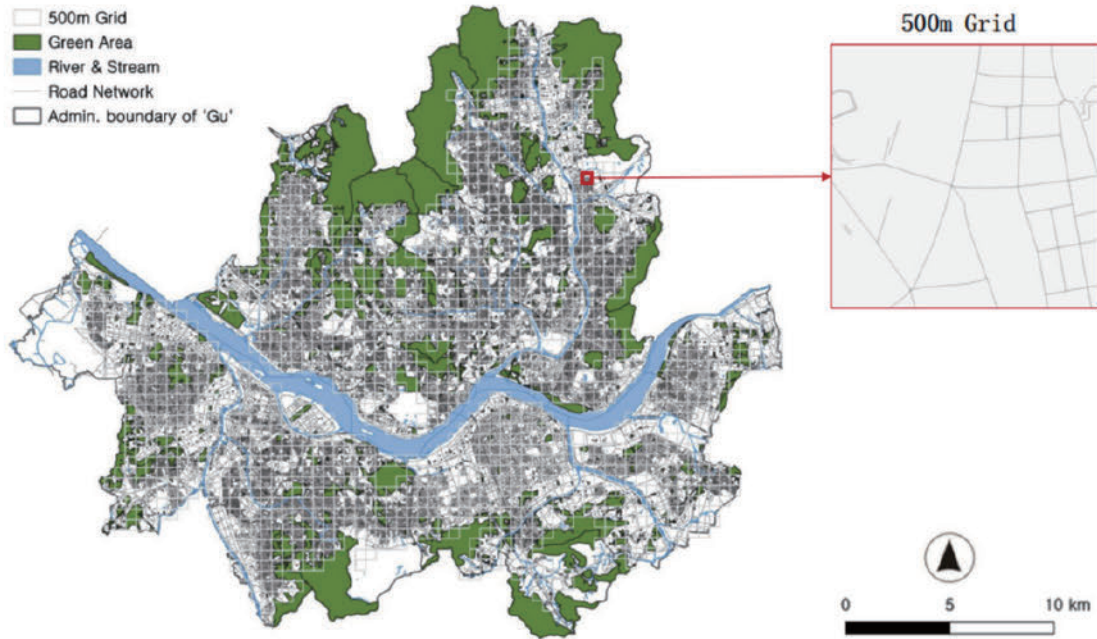


그림 1. 서울시 사례 지역과 500m 격자  
Figure 1. Case study area of Seoul and 500 m grid

## 2. 분석 과정

본 연구의 분석 과정은 <그림 2>와 같이 세 가지 주요 단계로 구성하였다. 먼저, POI 데이터를 전처리한 뒤 Word2vec 임베딩과 K-means 클러스터링을 적용하여 도시 기능구역을 유형화하였다. 다음으로, 도시 기능구역으로 유형화된 각 격자 단위로 교통사고 건수를 집계하여 종속변수를 구성하고, Street View Imagery(SVI)와 공간 빅데이터로부터 가시적 경관 지표를 포함한 도로 환경 지표를 구축하였다. 마지막으로, 기능구역별로 XGBoost 모형을 적합하고, SHAP 값을 활용해 변수 중요도를

산출하여 핵심 변수를 선별하였다. 이어 음이항 회귀모형을 적용하여 각 변수의 통계적 유의성을 검증하였다. 이러한 절차를 통해 변수 중요도 기반 선별과 계수 기반 추론을 결합한 혼합적 접근 프레임워크를 구현하였다.

## 3. 데이터 및 변수 구축

### 1) POI 데이터

POI 빅데이터는 실세계의 지리적 개체를 점(Point) 형태로 표현한 공간 빅데이터로, 공간적 현상과 인간 활동을 반영하는 정량적 분석의 주요한 자료로 활용된다(Zhu and Zhou, 2009; Wang et al., 2021). 본 연구에서는 도시 기능구역의 분류를 목적으로, 국내 대표적인 위치기반 서비스 플랫폼인 Kakao Maps POI를 활용하였다. 이는 다양한 유형의 장소 정보를 포괄적으로 수록하고 있으며, 그 정확성과 최신성 측면에서 도시공간 분석에 활용 가능한 유효한 데이터 소스로 평가된다(김선재·이수기, 2021). 본 연구에서는 카카오맵 API를 통해 2018년 6월부터 8월까지의 기간 동안 총 551,804건의 POI 데이터를 수집하였으며, 분류 체계는 카카오맵에서 제공하는 고유의 범주를 그대로 적용하였다. 이는 연구자 임의의 분류로 인한 해석의 자의성을 배제하기 위함이다. 다만, 카카오맵 POI는 '단독주택' 관련 정보가 상대적으로 미흡하다는 한계가 있어, 이를 보완하고자 통계청의 SGIS 데이터를 추가적으로 수집·가공하여 '단독주택' POI를 별도로 구축하였다. 수집된 POI의 유형 및 공간적 분포는 <표 1>과 <그림 3>에 각각 제시하였다.

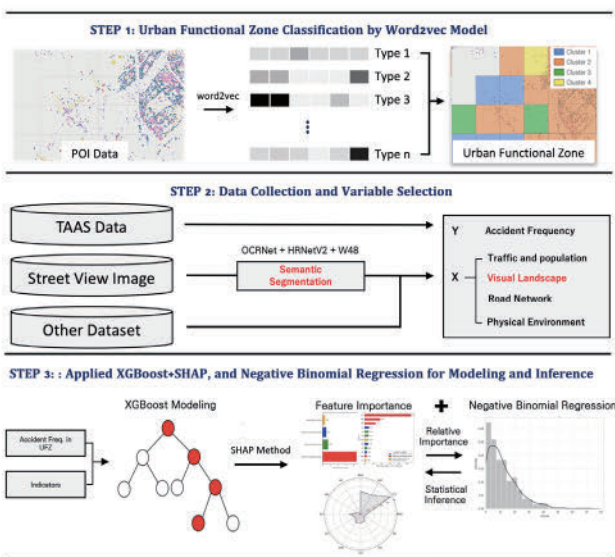


그림 2. 연구의 프레임워크  
Figure 2. Research framework

표 1. POI 분류

Table 1. POI classification

Classification	Detailed classification
단독 주택 Single-family housing	Single-family housing
아파트 Apartment	Apartment
다세대 주택 Multi-family housing	Villas, boarding houses, etc.
서비스/산업 Service/Industry	Household goods, medical equipment, etc.
금융/보험 Finance	Life Insurance, tax offices, credit guarantee institutions, etc.
음식점 Restaurant	Cafés, chicken restaurants, izakayas, etc.
의료/건강 Medical	Pharmacies, nursing care, orthopedic clinic, etc.
가정/생활 Lifestyle	Computer accessories, phone repair, hair salons, etc.
교육 Education	Elementary, middle and high schools, etc.
문화/예술 Culture/Arts	Exhibition halls, cinemas, theaters, etc.
사회/공공기관 Public institution	Senior welfare facilities, police boxes, community centers, etc.
스포츠/레저 Sports/Leisure	Gyms, golf ranges, billiard halls, etc.
언론/미디어 Media	Broadcasting stations, telecom companies, local cable TV, etc.
교통/수송 Transportation	Subway stations, express bus terminals, KTX stations, etc.
공원/하천 Park/River	Neighborhood parks, reservoirs, recreational areas, etc.



그림 3. POI 유형과 공간적 분포(서울시 강남구 예시)

Figure 3. POI types and spatial distribution (A case study of Gangnam-gu, Seoul, Korea)

## 2) 교통사고 데이터

본 연구에서는 500m 격자 단위의 자동차 교통사고 발생 빈도를 산출하기 위해 Traffic Accident Analysis System(TAAS) 자료를 활용하였다. TAAS는 경찰청, 보험사, 공제조합 등에서 수집된 다양한 유형의 교통사고 정보를 통합한 국가 단위의 데이터베이스로, 운전자의 인적 특성, 사고 심각도, 사고 발생지의 도로명 주소 등을 포함한다(황선근·이수기, 2021). 다만, 해당 자료는 사고 지점의 좌표 정보를 직접 제공하지 않으므로, 도로명 주소를 기반으로 지오코딩(Geocoding)을 수행하여 위·경도 좌표를 추정하였다.

또한 가시적 환경 요인의 효과를 명확히 파악하기 위해, 운전자의 시야 확보가 제한되는 야간 시간대 사고와 강우, 강설, 안개 등 악천후 상황에서 발생한 사고는 분석 대상에서 제외하였다. 전처리 후 총 21,606건의 사고 데이터를 최종 분석에 활용하였으며, 변환된 사고 지점 좌표를 기준으로 500m 격자와 공간 조인(Spatial Join)을 수행하여 격자별 교통사고 발생 건수를 집계하였다. 사고 지점의 공간적 분포는 <그림 4>에 제시하였다.

## 3) 분석 변수

본 연구에서는 설명변수는 가시적 경관 요인(Visual Landscape Factors), 교통 및 인구 요인(Traffic and Population Factors), 물리적 환경 요인(Physical Environment Factors), 도로 네트워크 요인(Road Network Factors)의 네 범주로 구성하였다(<표 2> 참조). 이 중 가시적 경관을 핵심 설명변수로, 나머지 세 범주는 노출 차이와 잠재적 교란을 보정하기 위한 통제변수(Controlled Variable)로 포함하였다. 분석에 활용된 세부 변수는 다음과 같다.

교통 및 인구 요인에는 도로 폭, 도로 폭 변동성, 교통혼잡비용, 인구 네 가지 변수로 구성하였다. 이들은 교통사고 발생에 유의미한 정(+)의 영향을 미치는 요인으로 다수의 선행연구에서 활

용된 바 있다(박준상, 2023; Yue, 2024). 특히, 교통혼잡비용 변수는 Annual Average Daily Traffic(AADT) 및 Volume-to-Capacity ratio(V/C)와 같은 정량적 교통지표 확보의 제약을 고려하여 대리 변수(Proxy Variable)로 활용되었다.

물리적 환경 요인은 도로 및 교통 인프라의 물리적 특성을 반영하며 교차로 수, 대중교통 시설 수, 평균 경사도를 선정하였다. 교차로 수는 운전자가 짧은 시간 내 정보를 인지하고 판단해야 하는 복잡한 주행 환경을 나타낸다(이지원·김태형, 2019), 대중교통 시설 수는 해당 지역의 통행 밀도 증가와 관련된 사고 위험을 반영한다. 경사도는 차량 제어와 주행 안정성에 영향을 미치는 물리적 조건으로(Fu et al., 2011) 선정되었다.

도로 네트워크 요인은 도로망의 구조적 특성을 반영하며 통과도(Betweenness)와 전환비(Diversion Ratio)를 사용하였다. 통과도는 특정 노드(교차로)의 중심성을 측정하며, 해당 지점의 교통 흐름 집중도와 사고 가능성을 설명하는 지표이다(Chang et al., 2022). 반면, 전환비는 복잡하고 불규칙한 도로 구조로 인해 운전자의 시야 전환 빈도가 높은 구간을 식별하는 지표로, 값이 높을수록 골목길 밀집도가 높고 사고 위험도 증가할 가능성이 있다(박은총·이수기, 2023; 성우석·이수기, 2025). 해당 변수는 QGIS 3.6.1의 sDNA 플러그인을 활용하여 산출하였고, 500m 격자 단위로 집계하였다. 변수에 대한 예시 그림은 <표 3>에 제시하였다.

마지막으로, 가시적 경관 요소는 운전자가 주행 중 시각적으로 인지하는 주변 환경을 정량화한 것으로, 본 연구에서는 녹지 지수(Green View Index, GVI), 시야 차단 지수(Visual Obstruction Index, VOI), 색상 복잡 지수(Color Entropy Index, CEI)를 주요 분석 항목으로 선정하였다. 녹지 지수는 파노라마 이미지에서 녹지 픽셀 비율을 의미하며, 시야 차단 지수는 시야 차단 객체(건물, 담장, 전봇대 등)와 교통 인프라 객체(차도, 보도, 신호기, 표지 등)의 상대적 비율을 기반으로 시야 확보 수준을 정량화한 지표이다. 마지막으로, 색상 복잡 지수는 도로 환경 내 색상 대비와 다채로움이 운전자의 시각적 피로와 정보 처리 부담에 미치는 영향을 반영하며, 박준상(2023)의 연구를 기반으로 도입되었다.

지표 산출 다음의 절차로 수행하였다. 먼저, 서울시 전역 도로망을 20m 간격으로 샘플링하여 총 95,842장의 Naver Street View 파노라마 이미지를 수집하였다. 의미론적 분할(Semantic Segmentation)은 영상의 각 픽셀을 사전 정의된 범주로 분류하는 기법으로, 본 연구에서는 OCRNet+HRNetV2+W48 딥러닝 모델을 적용하여 녹지(Vegetation), 건물(Building), 하늘(Sky), 차도(Road), 보행로(Sidewalk) 등 주요 객체의 픽셀 비율을 추출하였다. 해당 모델의 분할 성능은 Cityscape 데이터셋 기준 mean Intersection over Union (mIoU) 81.35를 기록하여(Yuan et al., 2020), 모델의 활용 타당성을 확보하였다. 이 결과를 바탕으로 GVI는 녹지 픽셀 비율로, VOI는 상기 식에 따라 계산하였다(식 (1), (2)).

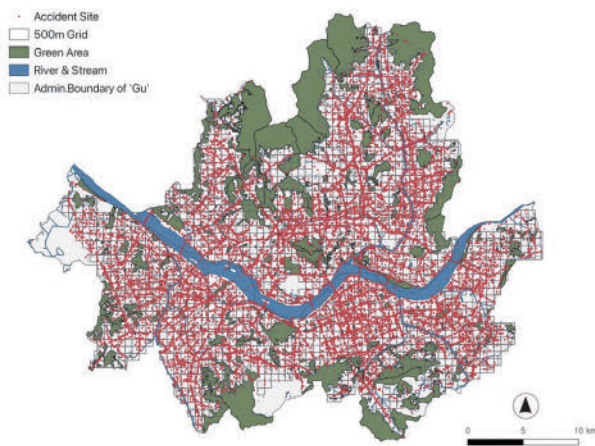


그림 4. 자동차 교통사고 발생지점의 공간적분포

Figure 4. The spatial distribution of traffic accident locations

표 2. 활용 변수 및 데이터 출처

Table 2. Description of variables and data source

Variables		Name	Definition	Data source	
종속변수 Dependent variable	자동차 교통사고 건수 Traffic accident frequency	-	Number of traffic accidents frequency in the grid	Traffic Accident Analysis System (2017~2019)	
	교통 및 인구요인 Traffic and population factor	도로 폭 Road width	RDW	Average road width within the grid	New Address Database (2018)
교통혼잡비용 Traffic congestion cost		TCC	Average traffic congestion cost within the grid (1 million won)	Real Estate Big Data Platform (2022)	
인구 Population		POP	Population within the grid (1,000 people)	New Address Database (2018)	
도로 폭 변동성 Road width variability		RWV	Standard deviation of road width within the grid		
독립변수 Independent variable	물리적 환경요인 Physical environment factors	교차로 수 Intersection	INT	Number of intersections within the grid	Seoul Open Data Plaza (2018)
		대중교통 시설 수 Subway and bus station	SBS	Number of subway and bus stations within the grid	Kakao Map API, Seoul Open Data Plaza (2018)
	경사도 Slope	SLP	Average slope within the grid	Real Estate Big Data Platform (2022)	
도로 네트워크 요인 Road network factors	통과도 Betweenness	BTW	Closeness of the road network within the grid	New Address Database (2018)	
	전환비 Diversion ratio	DVR	Diversion ratio of the road network within the grid		
가시적 경관 Visual landscape factors	녹지 지수 Greening view index	GVI	Quantifies the proportion of greenery in each panoramic image using semantic segmentation, aggregated within the grid	Naver Street View API (2018. 04-09.)	
	시야 차단 지수 Visual obstruction index	VOI	Calculated as the relative ratio of traffic infrastructure elements (e.g., roads, sidewalks, traffic signals, and traffic signs) to view-blocking elements (e.g., buildings, walls, fences, and utility poles), aggregated within the grid		
	색상 복잡 지수 Color entropy index	CEI	Assesses the diversity of colors in each panoramic image using k-means clustering and entropy analysis, aggregated within the grid		

표 3. 통합도 및 전환비 변수 예시

Table 3. Examples of betweenness and diversion ratio variables

Category	Traffic network 1	Traffic network 2
예시 Examples		
통합도 Betweenness	1.22	7.32
전환비 Diversion ratio	1.12	1.84

$$GVI = \frac{\sum_{i=1}^n Pixel_{Greenery_i}}{\sum_{i=1}^n Pixel_n} \tag{1}$$

$$VOI = \frac{\sum_{i=1}^n Pixel_{traffic\ sign_i + building + sidewalk_i + wall_i + fence_i + pole_i}}{\sum_{i=1}^n Pixel_n} \tag{2}$$

한편, 색상 복잡 지수는 이미지 내 각 픽셀의 RGB 값을 이용해 K-means 군집화를 수행한 후, 군집 분포의 엔트로피(식 (3))를 산출하였다(박준상, 2023). 군집 개수 K는 박근덕 외(2021)의 연구를 참고하여 16으로 설정하였다. 해당 변수는 0에서 1 사이의 값을 가지며, 0에 가까울수록 색상이 단조롭고 1에 가까울수록 색상 다양성이 높은 경관으로 해석된다. 각 변수의 처리 예시는 <그림 5>에 제시하였다.

$$CEI = \frac{\sum_{i=1}^n Pixel_i \times \log(Pixel_i)}{\log(16)} \quad (3)$$

#### 4. 분석 방법론

##### 1) 도시 기능구역 분류 방법

본 연구는 POI를 활용한 Word2vec 임베딩과 K-means 클러스터링 기법을 적용하여 도시 기능구역을 정량적으로 유형화하였다. 분석 절차는 이지원·김지영(2023)을 참고하여 구성하였다. Word2vec은 자연어 처리 분야에서 단어 간 의미적 유사성을 학습하는 대표적인 임베딩 기법으로, 단어를 고차원 벡터로 변환하여 문맥적 관계를 수치화한다. 본 연구에서는 이 기법을 POI 데

이터에 확장하여, 개별 POI 유형을 '단어', 인접한 POI의 조합을 '문맥'으로 간주함으로써 도시 내 활동의 의미적 관계를 벡터 공간에 반영하였다.

우선, POI를 기반으로 학습용 데이터셋을 구축하였다. 각 중심 POI에 대해 인접한  $k$ 개의 POI를 탐색하여 하나의 튜플을 구성하고, 중심-주변 POI 쌍의 물리적 거리 정보를 반영하기 위해 거리 기반 증강계수(Distance-Based Augmentation Factor)를 적용하였다. 해당 계수는 두 POI 간 거리가 가까울수록 더 큰 가중치를 부여하도록 설정하였으며, 구체적 정의는 Yan et al. (2017)을 참조한 식 (4)에 제시하였다. 예컨대, 특정 아파트 주변에 교육시설, 음식점, 의료시설이 공존하는 경우(아파트, 교육시설), (아파트, 음식점), (아파트, 의료시설)과 같은 튜플이 형성되며, 거리 가중을 통해 근접 관계가 학습에 더 크게 반영되도록 하였다.

$$\beta_{i,j} = \frac{1 + \frac{1}{|L|} \sum_{k=1}^{|L|} P_k}{1 + d^\alpha(l_i, l_j)} \quad (4)$$

여기서,  $\beta_{i,j}$ 는 POI  $l_i$ 와 POI  $l_j$  간 거리 기반 증강계수,  $d(l_i, l_j)$ 는 POI  $l_i$ 와 POI  $l_j$  간 거리,  $\alpha$ 는 거리 반비례 계수(본 연구에서는 1로 설정),  $|L|$ 는 전체 POI 개수,  $P_k$ 는 POI  $l_k$ 의 발생확

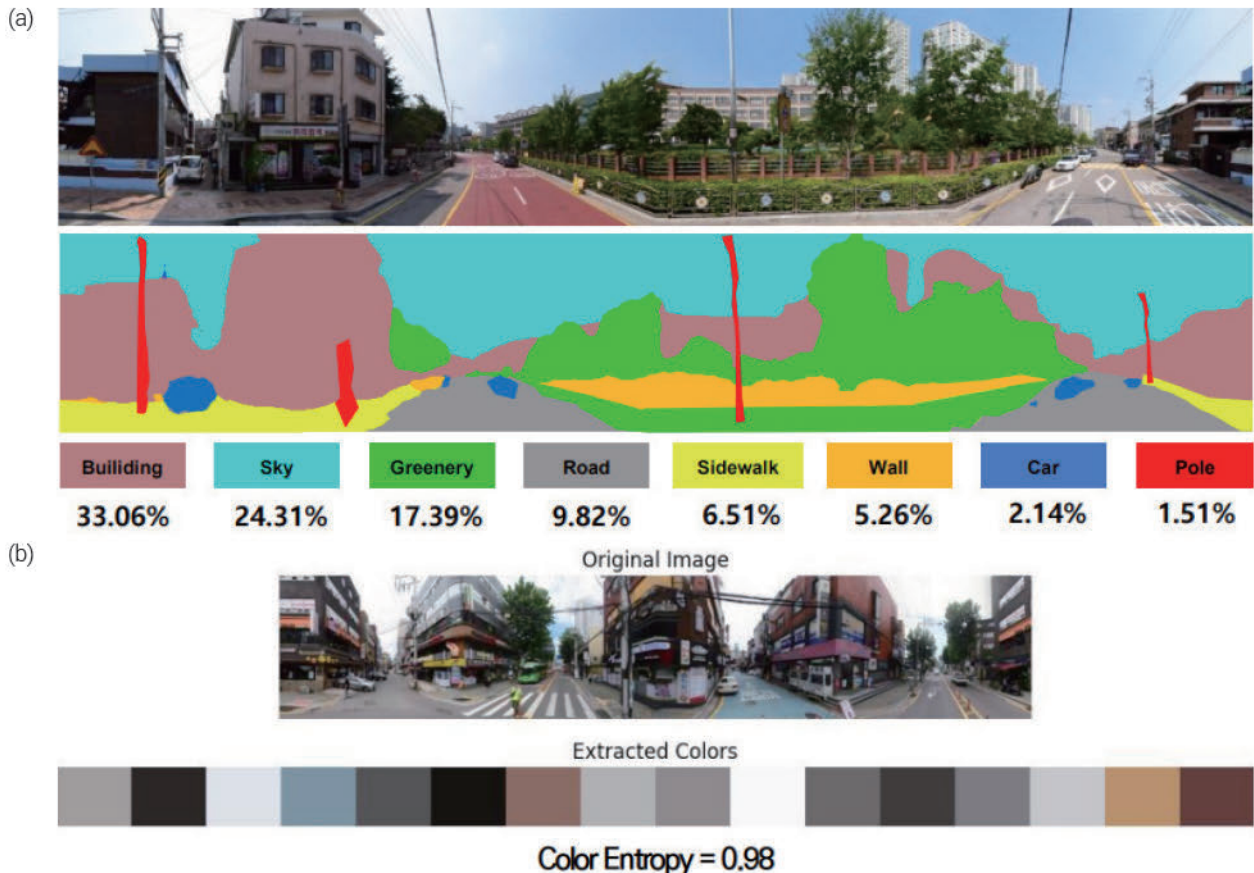


그림 5. 의미론적 분할 기법 (a) 및 color entropy (b) 예시  
Figure 5. Examples of semantic segmentation technique (a) and color entropy (b)

를로 정의된다.

임베딩 학습은 Word2vec의 Skip-gram 모델로 수행하였다. 중심 POI로 주변 POI를 예측하는 방식으로 파라미터를 추정하고, Softmax 정규화를 통해 출력 확률을 산출함으로써 POI 간 의미적 유사성을 반영한 벡터를 도출하였다. 도출된 벡터는 500m 격자 내 가중 평균하여 격자 대표 벡터를 구성하였다. 이를 입력값으로 K-means 클러스터링을 수행하여 k개의 도시 기능 구역을 분류하였고, 최적 클러스터는 Elbow 기법을 통해 결정하였다.

클러스터별 기능의 정량적 해석을 위해 특화계수(Location Quotient, LQ)를 산출하였다(식 (5)). 특화계수는 특정 POI 유형의 클러스터 내 밀도를 도시 전체 평균 밀도와 비교하여 상대적 집중도를 평가하는 지표로, 값이 1을 초과하면 해당 POI 유형이 해당 클러스터에서 상대적으로 집중되어 있음을 의미한다.

$$LQ_{i,c} = \frac{\sum_{j \in c} N_{i,j} / \sum_{j \in c} N_j}{\sum_{j \in C} N_{i,j} / \sum_{j \in C} N_j} \quad (5)$$

여기서,  $LQ_{i,c}$ 는 클러스터  $c$  내 특정 POI 유형  $i$ 의 밀도,  $N_{i,j}$ 는 격자  $j$ 에서 특정 POI 유형  $i$ 의 개수,  $N_j$ 는 격자  $j$  내 전체 POI 개수, 그리고  $c$ 는 특정 클러스터에 속하는 격자의 집합이다.

LQ 결과를 바탕으로 클러스터에서 특화된 주요 POI 유형을 도출하고 이를 근거로 대표 기능을 식별하였다. 명칭 부여의 객관성을 제고하기 위해, LQ로 추출된 주요 기능 항목을 입력으로 대규모 언어모델 기반(ChatGPT 4.5) 토픽 추출을 보조적으로 활용하여 클러스터 명칭을 부여하였다. 최종적으로, 도출된 도시 기능구역은 위성영상 자료와의 대조를 통해 공간적 타당성을 확인하였다.

## 2) XGBoost+SHAP

XGBoost는 그래디언트 부스팅(Gradient Boosting) 기반의 대표적인 기계학습 알고리즘으로, 높은 예측 성능과 과적합 방지에서 강점을 보인다. 다만 모델 구조의 복잡성으로 인해 해석 가능성이 제한된다. 이에 본 연구에서는 SHAP 기법을 결합하여 변수별 기여도를 정량적으로 평가하였다. SHAP는 각 변수의 예측 기여도를 다음과 같이 분해한다(식 (6)).

$$f(x) = \Phi_0 + \sum_{i=1}^M \Phi_i \quad (6)$$

여기서,  $\Phi_0$ 은 전체 평균 예측값,  $\Phi_i$ 는  $i$ 번째 변수의 기여도(Shapley 값)이다. 이를 통해 각 변수의 상대적 영향력을 직관적으로 시각화할 수 있다.

## 3) 음이항 회귀모형

음이항 회귀모형은 포아송 회귀의 확장으로, 과대산포가 존재하는 이산형 빈도자료에 적합하다. 해당 모형은 식 (7)과 같은 로그링크 함수로 구성된다.

$$\log(\mu_i) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} \quad (7)$$

여기서,  $X_k$ 는 시각적 경관 요소를 포함하는 독립변수, 계수  $\beta_k$ 는 계수로 크기와 방향성으로 사고 발생에 미치는 영향을 해석한다.

## IV. 분석 결과

### 1. 도시 기능구역 분류

#### 1) POI 임베딩 및 클러스터 수 결정

총 15개 POI 유형을 대상으로 각 POI를 중심으로 인접한 10개 POI를 묶어 거리 기반 중강계수를 적용한 결과, 총 6,311,320개의 학습 튜플을 확보하였다. 이후 Word2Vec의 Skip-gram으로 벡터 차원 70(이지원·김지영, 2023)과 epoch 1,000으로 설정하여 임베딩을 수행하였다. 도출된 POI 유형별 임베딩은 <표 4>에 제시하였다. 이후 각 500m 격자의 POI 임베딩 가중 평균으로 대표 벡터를 구성해 K-means로 도시 기능구역을 유형화하였으며, Elbow 기법 결과 k=4가 최적 클러스터 수로 선정되었다(<그림 6> 참조).

#### 2) 클러스터별 도시기능 유형화

클러스터별 기능을 정량적으로 해석하기 위해 POI 유형별 LQ를 산출하였다(<표 5> 참조). 그 결과, 다수 클러스터에서 여러 POI 유형의 LQ가 1을 상회하여, 특정 단일 기능의 독점적 지배

표 4. Word2vec 기반 POI 임베딩 결과

Table 4. Word2vec-based POI embedding results

POI type	Embedding results (70 dimensions)
단독주택 Single-family housing	[0.26320,0.01847,-0.00683,...,0.01189]
아파트 Apartment	[0.36258,-0.16685,0.16848,...,-0.19885]
다세대 주택 Multi-family housing	[0.56258,0.96285,-0.16918,...,-0.15695]
	⋮
언론/미디어 Media	[0.44558,-0.39424,0.15048,...,0.56885]
교통/수송 Transportation	[0.21871,-0.01533,0.34539,...,0.51906]
공원/하천 Park/River	[0.29963,-0.00759,0.01815, ...,0.08598]

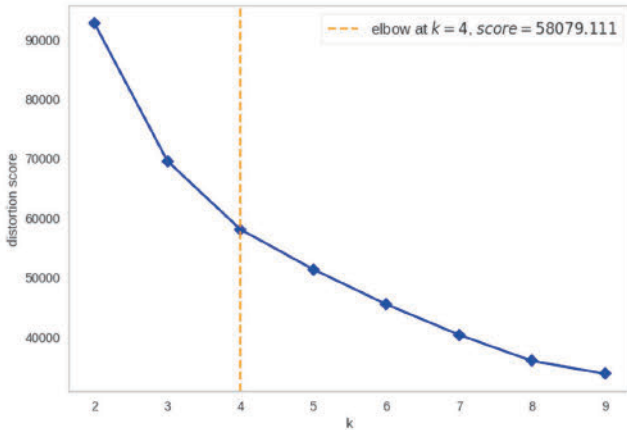


그림 6. Elbow 결과  
Figure 6. Elbow method result

표 5. POI 유형별 특화계수 산출 결과

Table 5. Results of POI enrichment factor by type

Category	POI enrichment factor			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
단독 주택 Single-family housing	0.232	0.356	0.088	<b>4.118 (1st)</b>
아파트 Apartment	<b>6.456 (1st)</b>	0.807	0.519	0.593
다세대 주택 Multi-family housing	0.960	<b>1.017 (3rd)</b>	0.264	0.956 (2nd)
서비스/산업 Service/Industry	0.511	<b>1.132 (2nd)</b>	0.775	0.525
금융/보험 Finance/Insurance	0.222	<b>1.207 (1st)</b>	<b>1.338 (3rd)</b>	0.240
음식점 Restaurant	0.492	1.009	0.471	0.645
의료/건강 Medical/Health	0.750	1.003	0.273	0.615
가정/생활 Lifestyle	0.636	1.008	0.385	0.743
교육/학문 Education/Academic	<b>2.388 (3rd)</b>	0.986	0.387	<b>0.751 (3rd)</b>
문화/예술 Culture/Arts	0.580	1.004	0.530	0.480
사회/공공기관 Public Institution	1.482	1.003	1.238	0.729
스포츠/레저 Sports/Leisure	1.334	1.015	1.020	0.627
언론/미디어 Media	0.399	1.016	0.320	0.459
교통/수송 Transportation	1.322	0.997	<b>9.352 (2nd)</b>	0.690
공원/하천 Park/River	<b>3.863 (2nd)</b>	0.772	<b>20.330 (1st)</b>	0.785

보다는 복합 기능의 공존이 관찰되었다. 이는 서울시의 혼합적 기능배치 특성에도 부합한다. LQ 구성을 근거로 대표 기능을 식별하고, 대규모 언어모형(LLM) 기반 토픽 추출과 연구진 간 검토를 통해 객관성과 해석 타당성을 고려한 최종 명칭을 확정하였다(〈표 6〉 참조).

클러스터 1은 아파트(LQ=6.456), 공원 및 하천(3.863), 교육 및 학문(2.388) 유형에서 가장 높은 특화계수를 나타냈다. 이는 고층 공동주택 밀집 지역에 쾌적한 녹지와 교육시설이 복합된 고밀 주거지역의 전형적인 공간구조를 반영한다. 이에 따라 클러스터 1은 ‘고밀 주거 및 교육 복합 기능구역(High-density Residential & Educational Mixed-use Zone)’으로 명명하였다. 클러스터 2는 금융 및 보험(1.207), 서비스 및 산업(1.132), 다세대 주택(1.017), 음식점(1.009) 등 다양한 일상생활 및 상업기반 POI가 밀집된 혼합 기능구조를 보였다. 이에 본 클러스터는 ‘생활 복합 기능구역(Urban Mixed-use Activity Zone)’으로 명명하였다. 클러스터 3은 교통 인프라와 여가·녹지 기능이 동시에 밀집된 공간 구조를 보였다. 이에 따라 본 구역은 ‘교통 및 여가 복합 기능구역(Transit and Leisure Mixed-use Zone)’으로 명명하였다. 마지막으로, 클러스터 4는 단독주택(4.118)이 뚜렷하게 특화되고, 교육 및 학문(0.751), 가정/생활(0.743)이 보조적으로 분포하여 전통적인 저층 주거지역의 구조를 유지하면서도 지역 기반의 생활 편의시설이 함께 분포된 저밀도 주거지의 특성을 시사한다. 이에 본 클러스터는 ‘저밀 주거 및 생활 복합 기능구역(Low-density Residential & Local Living Zone)’으로 명명되었다.

### 3) 기능구역 분류의 공간적 타당성 검토

본 연구는 서로 다른 클러스터 경계부를 중심으로 위성영상 대조를 통해 군집 결과와 현지 공간 맥락의 정합성을 검토하였다(그림 7). 고밀 주거 및 교육 복합 기능구역(Cluster 1)은 ‘코오롱 아파트’와 ‘서울 우암초등학교’ 등이 위치한 지역으로, 위성영상에서도 대단위 아파트 단지와 교육시설, 인접한 녹지공간이 복합적으로 분포하고 있어 주거-교육 중심지로서의 공간 특성이 뚜렷하게 확인되었다. 생활 복합 기능구역(Cluster 2)은 ‘LG전자’, ‘KT 양재지사’, ‘양재1동 주민센터’, ‘다세대 주택’ 등이 혼재된 지역으로, 상업 및 공공 서비스 기능이 밀집하고 다양한 용도의 건축물과 생활 기반 시설이 복합된 도시 일상생활 중심지의 특성을 보였다. 교통 및 여가 복합 기능구역(Cluster 3)은 ‘경부고속도로’, ‘전기차 충전소’, ‘어린이 교통공원’ 등이 위치한 지역으로, 교통 인프라의 밀집과 여가 공간의 혼재가 동시에 관찰되었다. 저밀 주거 및 생활 복합 기능구역(Cluster 4)은 단독주택과 대학교, 근린 생활 편의시설 등이 분포하는 저밀도 주거지로, 위성영상에서도 지역 커뮤니티 중심의 공간 구성 등이 명확히 나타났다. 종합적으로, 제시된 기능구역은 도시의 실질적 구조적 특성을 효과적으로 반영하고 있음을 확인하였다.

표 6. 도시 기능구역 클러스터 분류 결과

Table 6. Results of urban functional zone cluster classification

No.	Urban function zone	Distinctive POI types	Key characteristics
1	고밀 주거 및 교육 복합 기능구역 High-density residential & educational mixed-use zone	Apartments (Enrichment factor: 6.456), Education & academic (2.388), Green Spaces/Waterways (3.863)	A densely developed area combining high-rise apartments, green environments, and educational facilities.
2	생활 복합 기능구역 Urban mixed-use activity zone	Finance & insurance (1.207), Services & industry (1.132), Multi-unit dwelling (1.017), Restaurant (1.009)	A multifunctional urban space where daily life infrastructure, commercial, and service-related POIs are concentrated.
3	교통 및 여가 복합 기능구역 Transit and leisure mixed-use zone	Green spaces/Waterways (20.330), Transportation & transit (9.352), Finance (1.338)	A spatially mixed zone where large-scale transport infrastructure coexists with green and leisure elements.
4	저밀 주거 및 생활 복합 기능구역 Low-density residential & local living zone	Detached housing (4.118), Home & living (0.743), Education & academia (0.751)	A low-density residential district centered on single-family homes and community-level facilities.

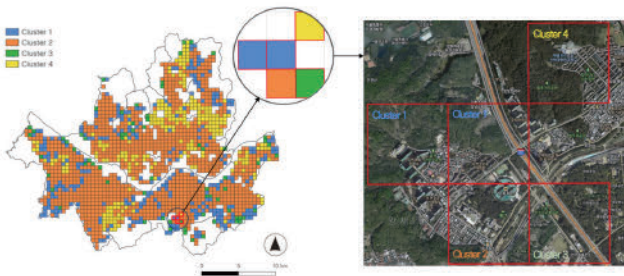


그림 7. 도시 기능구역 클러스터 및 검증  
Figure 7. Urban functional zone clusters and validation

## 2. 기술 통계분석 결과

도시 기능구역 간 이질성을 확인하기 위하여 클러스터별 기술 통계를 산출하였다(표 7 참조). 그 결과, 대부분의 변수에서 분포가 뚜렷이 구분되었으며, ANOVA 검정에서는 모든 변수의 클러스터 간 평균 차이가 통계적으로 유의한 것으로 나타났다. 이는 기능구역별로 교통사고 발생 빈도와 환경적 특성의 실질적 차이를 시사한다.

종속변수인 자동차 교통사고 건수의 전체 평균은 11,858건, 표준편차 10,468로, 분산이 평균을 상회하는 과분산 특성이 확인되었다. 이러한 분포 특성은 각 기능구역 수준에서도 유사하게 관찰되었으며(그림 8 참조), 이에 따라 본 연구는 음이항 회귀모형을 적용하고자 하였다.

교통 및 인구 요인에서, 평균 도로 폭은 교통 및 여가 복합 기능구역(Cluster 3)이 17.78m로 가장 넓은 반면, 저밀 주거 중심구역(Cluster 4)은 5.92m로 가장 협소하였다. 교통혼잡비용은 상업 및 업무 기능이 밀집한 Cluster 2에서 평균 12.24로 최대치를 보이며, 높은 통행 수요를 반영한다. 반면, Cluster 3은 광역 교통 기반 인프라가 존재함에도 인구 밀도가 평균 0.396으로 낮게 나타났다.

물리적 환경 요인에서도 유사한 경향이 확인되었다. Cluster 2는 교차로 수(1,775)와 대중교통 시설 수(6,972) 모두에서 가장 높은 값을 기록하여 도시 활동이 집약된 중심지의 특성을 보였다. 반대로 교통 및 여가 복합 기능구역(Cluster 3)은 해당 지표에서 모두 최저 수준이었는데, 이는 광역 교통 인프라(예: 고속도로, 대형 간선도로)가 주를 이루고, 대중교통 및 교차로, 환승거점 등과 같은 미시적 네트워크의 밀도는 상대적으로 낮은 자동차 중심의 교통 특성이 반영된 결과로 해석된다.

도로 네트워크 요인의 경우, Cluster 4가 통과도(3,601)와 전환비(1,606)에서 가장 높았다. 이는 구역 내 주요 도로의 대외 연결성이 큰 구조임을 시사한다. 반면, Cluster 1과 3은 통과도 지표가 상대적으로 낮게 나타났다. Cluster 1의 경우, 밀집된 주거 및 교육시설에 따라 외부와의 도로 연결성이 일부 제한된 구조로 해석될 수 있으며, Cluster 3의 경우에는 특정 주요 간선로를 통한 이동이 주를 이루는 공간 특성이 반영된 것으로 볼 수 있다.

가시적 경관 요인에서는 GVI가 클러스터 1(26,546)과 클러스터 3(25,618)에서 높게 나타나 고층 주거지-녹지/여가 공간의 결합을 시사한다. VOI는 클러스터 4에서 평균 4.485로 최대치를 보여, 협소한 가로와 높은 건축밀도에 따른 시야 제한 가능성을 시사한다. CEI는 전반적으로 0.918-0.935범위로 유사하나, 클러스터 3의 표준편차(0.071)가 가장 커 시각적 다양성의 공간적 편차가 큰 것으로 나타났다. 마지막으로 다중공선성(Variance Inflation Factor, VIF)을 검토한 결과, 클러스터별 모든 변수의 VIF 값이 5 이하로 나타나, 변수 간 독립성이 확보되었음을 확인하였다.

## 3. 도시 기능구역별 교통사고 패턴

도시 기능구역 간 사고 수준을 비교하기 위해 교통사고 빈도와 평균 사고 빈도( $\mu$ ), 표준편차(std)를 기준으로 다섯 수준으로 범주화하여 공간적 패턴을 도출하였다.

표 7. 기술통계분석 결과  
Table 7. Descriptive analysis

Category	Variables	Descriptive statistics						ANOVA test		
		Obs.	Cluster	Mean	Std.Dev.	Min.	Max.	F	p	
종속변수 Dependent variable	자동차 교통사고 건수 Car accident count	1,714	전체	11.858	10.468	0.000	72.000			
		284	1	5.616	6.189	0.000	39.000			
		1,109	2	14.378	10.788	0.000	72.000			
		69	3	5.899	8.907	0.000	37.000			
		252	4	9.433	9.036	0.000	50.000			
		<b>ANOVA</b>							75.82***	<0.001
독립변수 Independent variables	도로 폭 Road width	1,714	전체	10.531	6.768	1.750	49.800			
		284	1	14.897	8.480	3.186	49.800			
		1,109	2	10.055	5.697	2.942	45.400			
		69	3	17.780	10.412	1.750	44.000			
		252	4	5.915	1.719	2.644	12.727			
		<b>ANOVA</b>							127.85***	<0.001
	교통 및 인구요인 Traffic and population factors	도로 폭 변동성 Road width variability	1,714	전체	7.615	3.805	0.000	32.025		
			284	1	7.347	3.661	0.000	24.404		
			1,109	2	8.183	3.674	0.000	32.025		
			69	3	6.958	6.583	0.000	22.490		
			252	4	5.597	2.580	1.237	15.676		
<b>ANOVA</b>							34.99***	<0.001		
독립변수 Independent variables	교통혼잡비용 Traffic congestion cost	1,714	전체	10.362	9.572	0.009	64.954			
		284	1	6.588	7.006	0.009	44.127			
		1,109	2	12.240	9.787	0.009	57.669			
		69	3	8.367	12.609	0.009	64.954			
		252	4	6.896	7.697	0.009	50.195			
<b>ANOVA</b>							44.05***	<0.001		
독립변수 Independent variables	인구 Population	1,714	전체	5.167	3.364	0.005	14.916			
		284	1	4.736	3.239	0.005	13.971			
		1,109	2	5.406	3.279	0.005	14.916			
		69	3	0.396	0.740	0.005	5.105			
		252	4	5.908	3.239	0.039	13.118			
<b>ANOVA</b>							59.27***	<0.001		
물리적 환경요인 Physical environment factors	교차로 수 Intersection	1,714	전체	1.569	1.459	0.000	9.000			
		284	1	1.278	1.420	0.000	9.000			
		1,109	2	1.775	1.485	0.000	8.000			
		69	3	0.594	0.880	0.000	3.000			
		252	4	1.258	1.294	0.000	6.000			
<b>ANOVA</b>							26.29***	<0.001		

다음 페이지에 계속(Continue on next page)

Category	Variables	Descriptive statistics						ANOVA test				
		Obs.	Cluster	Mean	Std.Dev.	Min.	Max.	F	p			
독립변수 Independent variables	물리적 환경요인 Physical environment factors	1,714	전체	6.222	4.498	0.000	28.000	62.34***	<0.001			
		284	1	4.299	3.840	0.000	20.000					
		1,109	2	6.972	4.405	0.000	28.000					
		Subway and bus station	69	3	1.217	1.909	0.000			7.000		
		252	4	6.460	4.601	0.000	26.000					
		ANOVA										
		1,714	전체	12.818	7.655	0.400	60.786			11.51***	<0.001	
		284	1	14.785	7.627	0.400	59.187					
		1,109	2	12.112	7.837	0.400	60.786					
		Slope	69	3	14.941	6.096	3.895					35.607
	252	4	13.129	6.720	0.400	35.084						
	ANOVA											
	도로 네트워크 요인 Road network factors	1,714	전체	2.713	1.174	0.080	8.629	107.75***	<0.001			
		284	1	2.075	1.236	0.080	6.918					
		1,109	2	2.734	1.016	0.080	6.463					
		통과도 Betweenness	69	3	1.778	1.297	0.080					5.262
		252	4	3.601	1.083	1.579	8.629					
		ANOVA										
		전환비 Diversion ratio	1,714	전체	1.529	0.137	1.149			1.948	88.42***	<0.001
			284	1	1.446	0.163	1.149			1.945		
			1,109	2	1.539	0.121	1.154			1.948		
			69	3	1.423	0.150	1.149			1.789		
	252		4	1.606	0.100	1.302	1.913					
	ANOVA											
녹지 지수 Greening view index	1,714		전체	17.593	11.476	0.088	69.572	103.34***	<0.001			
	284		1	26.546	12.988	2.542	68.248					
	1,109		2	15.567	9.772	0.659	69.572					
	69		3	25.618	14.200	0.088	62.823					
	252	4	14.222	9.724	3.364	67.130						
	ANOVA											
	가시적 경관 Visual landscape factors	1,714	전체	3.687	1.698	0.247	14.840			90.10***	<0.001	
		284	1	2.464	1.507	0.384	8.519					
		1,109	2	3.872	1.554	0.247	13.073					
		시야 차단 지수 Visual obstruction index	69	3	2.829	2.138	0.384					14.840
252		4	4.485	1.593	0.316	13.129						
ANOVA												
색상 복잡 지수 Color entropy index		1,714	전체	0.934	0.018	0.372	0.965	18.32***	<0.001			
		284	1	0.933	0.015	0.849	0.965					
		1,109	2	0.935	0.010	0.763	0.965					
		69	3	0.918	0.071	0.372	0.955					
	252	4	0.934	0.008	0.898	0.948						
	ANOVA											

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1

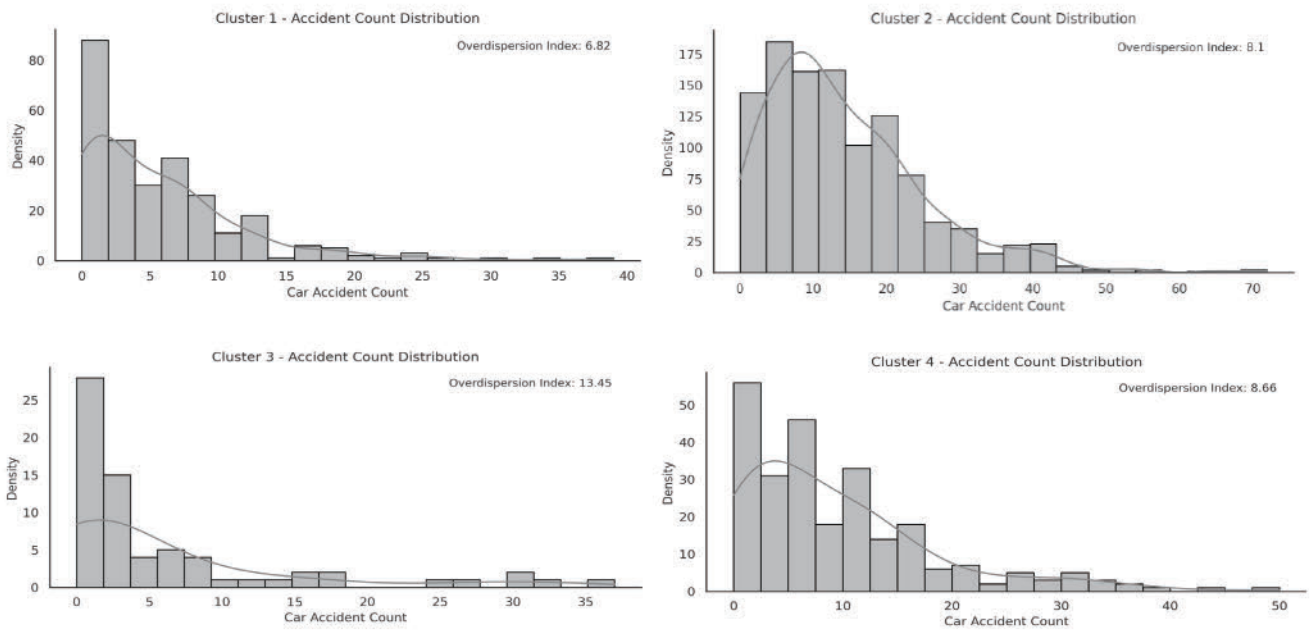


그림 8. 클러스터별 교통사고 빈도 분포

Figure 8. Distribution of traffic accident frequencies by urban functional zones

- 1) 높음(High):  $\text{Accident Count} \geq \mu + \text{std}$
  - 2) 약간 높음(Sub-High):  $\mu + 0.5\text{std} \leq \text{Accident Count} \leq \mu + \text{std}$
  - 3) 중간(Medium):  $\mu - 0.5\text{std} \leq \text{Accident Count} \leq \mu + 0.5\text{std}$
  - 4) 약간 낮음(Sub-Low):  $\mu - \text{std} \leq \text{Accident Count} \leq \mu - 0.5\text{std}$
  - 5) 낮음(Low):  $\text{Accident Count} < \mu - \text{std}$
- 전역적 사고 분포를 시각화한 결과(그림 9(a)), 강남, 여의도,

한양도성 등 주요 도심부에 집중되고 외곽으로 갈수록 낮아지는 경향을 보였다. 기능구역별 비교를 위해, 사고 수준을 범주화하여 박스플롯으로 제시하고(그림 9(b)), 수준별 면적 비율을 산출하였다(그림 9(c)). 그 결과, 상업 및 공공서비스 중심 지역(Cluster 2)과 저밀도 주거 및 생활 밀착형 지역(Cluster 4)에서 상대적으로 높은 사고 수준이 관찰되었다. 특히, 클러스터 2에서는 Medium 44.3%, High 18.5%, 클러스터 4에서도 Medium 38.1%로 높았다. 이는 상업 활동이 활발하거나 생활도어가 밀집된 지역에서 교통사고 위험이 상대적으로 높음을 시사한다. 반

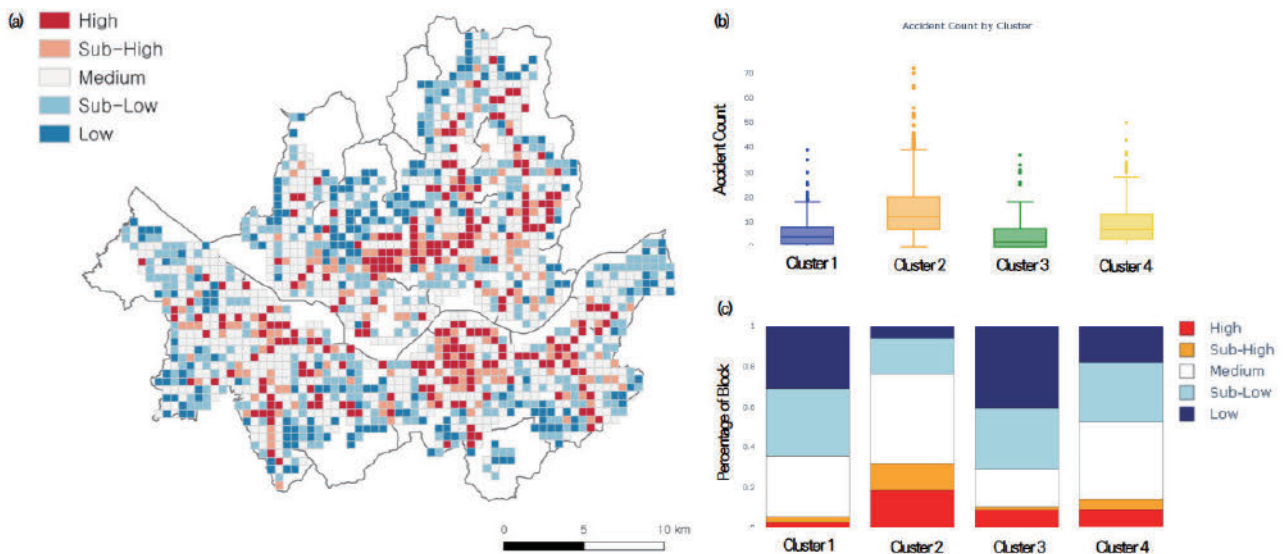


그림 9. 교통사고의 공간적 패턴과 도시 기능구역별 교통사고 빈도의 box plots 및 강도 비율

Figure 9. Spatial patterns of traffic accidents and box plots of accident frequency ratios by urban functional zone

면, 고밀 주거 및 교육 중심 지역(Cluster 1)과 교통 및 여가 중심 지역(Cluster 3)은 전반적으로 낮은 수준의 분포를 보였고, 두 구역의 분포 양상 또한 유사하였다. 요약하면, 교통사고 발생 패턴은 공간·기능 맥락에 따라 뚜렷이 분화된다. 상업·생활도로 중심 구역에서는 위험이 상대적으로 높고, 주거·교육·교통·여가 중심 구역에서는 비교적 안정적인 양상이 확인되었다.

#### 4. 변수 중요도 및 음이항 회귀분석

##### 1) XGBoost 모형 평가

교통사고 발생에 영향을 미치는 변수의 상대적 중요도를 평가하기 위해 XGBoost 모형을 구축하였다. 모형의 일반화 성능 확보를 위해 Bayesian Optimization을 통한 하이퍼파라미터 최적화와 5-fold 교차 검증을 적용하였다. 분석 결과, 도시 기능구역별 XGBoost 모형의 결정계수( $R^2$ )는 0.578~0.716 범위로 나타났다(〈표 8〉 참조), 이는 박준상(2023)의 서울시 전체 분석 결과( $R^2=0.52$ )보다 높은 수준이다. 이러한 결과는 본 연구의 모형이 일정 수준 이상의 예측력을 갖추었음을 시사하며, 후속 SHAP 분석 결과에 대한 충분한 설명력을 확보한 것으로 판단되었다.

표 8. XGBoost 모형 평가

Table 8. Evaluation of XGBoost model

Cluster	Model performance		
	$R^2$	RMSE	MAE
1	0.578	4.418	2.956
2	0.604	6.396	4.850
3	0.703	4.898	2.972
4	0.716	4.893	3.713

##### 2) 변수 중요도 분석

XGBoost 모형의 SHAP 값을 활용하여 클러스터별 변수 기여도를 산출하고, 레이더 차트와 막대그래프로 시각화하였다(〈그림 10〉 참조). 레이더 차트는 기능구역별 변수 중요도의 전반적 분포와 형태 차이를 직관적으로 보여주며, 구역 유형에 따라 사고 예측에 기여하는 요인 구성이 구조적으로 상이함을 시사한다. 전반적으로 '교통 및 인구' 범주의 중요도가 높게 나타났으나, 일부 가시적 경관 변수도 평균을 상회하는 중요도를 기록한 점이 주목할 만하다.

구체적으로 고밀 주거·교육 중심 지역(Cluster 1)에서는 교통 혼잡비용(TCC: 2.74)이 가장 높은 기여도를 보였으며, 색상 복잡 지수(CEI: 0.23)와 녹지 지수(GVI: 0.21) 역시 전체 평균(0.18)을 상회하였다. 생활 복합 지역(Cluster 2)에서는 녹지 지수(GVI: 1.50)가 평균 대비 8.3배로 가장 높았고, 도로 폭 변동성(RWV: 1.59)과 교통혼잡비용(TCC: 4.88)도 높은 중요도를 나타냈다. 교통·여가 복합 지역(Cluster 3)의 경우 교통혼잡비용(TCC: 2.88), 도로 폭 변동성(RWV: 1.76), 경사도(SLP: 1.16)가 주요 변수로 도출되었으며, 시야 차단 지수(VOI: 0.54)는 평균 대비 약 3배 높은 값을 기록하였다. 마지막으로, 저밀 주거·생활 복합 지역(Cluster 4)에서는 교통혼잡비용(TCC: 3.33), 도로 폭 변동성(RWV: 1.98), 인구(POP: 0.96) 외에 녹지 지수(GVI: 0.66)가 평균 대비 3.7배 높은 수준으로 나타났다. 이러한 결과는 도시 기능구역별 간 사고 예측에 기여하는 요인 구성이 상이하며, 일부 가시적 경관 요소가 특정 공간 맥락에서 높은 설명력을 가질 수 있음을 시사한다.

##### 3) 음이항 회귀분석 결과(통계적 검증)

상기 XGBoost-SHAP 분석을 통해 변수 중요도가 평균 이상으로 나타난 주요 변수들을 중심으로, 기능구역별 음이항 회귀모

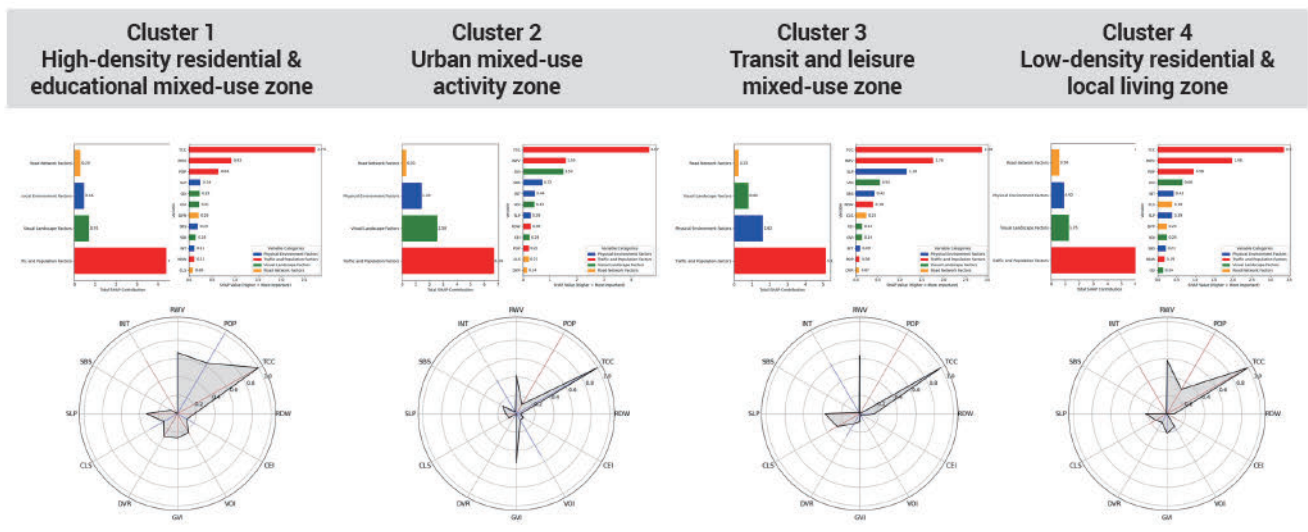


그림 10. 도시 기능구역별 특성 중요도 결과  
Figure 10. Results of urban functional zone feature importance

형을 적용하여 통계적 유의성을 검증하였다(〈표 9〉 참조). 분석 결과는 다음과 같다.

모든 클러스터에서 교통 및 인구 특성 변수는 사고 발생과 일관된 정(+)의 연관성을 보였다. 이는 Chang et al.(2022) 등 기존 선행연구와도 부합하는 결과로, 전통적인 물리·인구 요인이 사고 발생에 핵심적인 영향을 미침을 확인할 수 있다. 반면, 가시적 경관 요소의 효과는 구역별로 상이한 통계적 양상을 보였다.

고밀 주거 및 교육 복합 기능구역(Cluster 1)에서는 색상 복잡 지수(CEI)가 유의한 부(-)의 효과를 보였다. 반면, 녹지 지수(GVI)는 SHAP 분석에서 높은 중요도를 나타냈으나, 통계적으로 유의한 연관성이 확인되지 않았다. 이는 고밀 환경에서 녹지의 독립적 영향력이 주변 교차로 밀도, 도로 혼잡 등과 상호작용하여 상쇄될 가능성을 시사한다. 생활 복합 기능구역(Cluster 2)에서는 녹지 지수(GVI)가 유의한 부(-)의 효과를 나타내, 녹지가 시각적 완충 혹은 속도 저감 메커니즘과 연동되어 실제 사고 발생 저감에 기여할 가능성을 시사한다. 교통 및 여가 복합 기능구역(Cluster 3)에서는 시야 차단 지수(VOI)가 유의한 양(+)의 효과를 보여, 대형 교통 인프라와 여가시설이 혼재된 구역에서 물리적 시야 방해 요소가 사고 발생의 결정적 요인이 될 수 있음을 시사한다.

마지막으로, 저밀 주거 및 생활 복합 기능구역(Cluster 4)에서는 녹지 지수(GVI)가 평균 이상의 중요도를 보이며, 통계적으로도 사고 위험 증가에 유의한 양(+)의 효과를 나타냈다. 이는 좁은 도로 및 무계획적 녹지 배치가 시야 확보를 저해하여 오히려 교통 사고 위험을 높이는 원인으로 작용할 가능성을 시사한다. 요컨대, 각 도시 기능구역 간 SHAP 분석에서 중요도가 높게 나타난 가시적 경관 요소들이 회귀분석에서도 통계적으로 유의하게 작동한 경우가 다수 관찰되었다. 다만 일부 변수는 중요도는 높았으나 통계적으로는 유의하지 않은 사례도 확인되었다. 이러한 결과는, 도시의 공간적 맥락과 물리적 환경에 따라 동일 변수의 효과가 달라질 수 있음을 보여주며, 이에 대한 심층적 해석과 정책적 논의는 5장 토의에서 상세히 다룬다.

## V. 토의 및 결론

### 1. 토의

본 연구는 도시 기능구역별로 가시적 경관 요소가 자동차 교통 사고에 미치는 맥락적 차별성을 실증적으로 규명하였다. 분석 결과와 선행연구를 바탕으로, 다음과 같은 주요 논의점을 제시한다.

표 9. 음이항 회귀분석 결과

Table 9. Negative binomial regression results

Category	Variables	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
		Coef.	p	Coef.	p	Coef.	p	Coef.	p
Traffic and population factors	RDW	-0.815	0.194	-0.309	0.480	1.024	0.392	2.972	0.257
	TCC	5.717***	0.000	2.393***	0.000	3.314***	0.001	2.722***	0.000
	POP	1.482***	0.000	0.294*	0.083	12.062***	0.000	1.115**	0.033
	RWV	3.450***	0.000	2.134***	0.000	2.200**	0.022	4.840***	0.000
Physical environment factors	INT	0.051	0.915	0.489**	0.024	1.386	0.518	0.814	0.152
	SBS	0.158	0.797	0.410*	0.082	-1.390	0.656	-0.035	0.853
	SLP	-0.021	0.955	-0.184	0.466	8.375***	0.000	-0.621	0.293
Road network factors	CLS	1.433*	0.052	0.025	0.937	-0.050	0.979	-0.461	0.614
	DVR	0.717*	0.097	-0.175	0.551	0.413	0.720	0.508	0.463
Visual landscape factors	녹지 지수 (GVI)	0.070	0.904	-1.406***	0.000	1.317	0.314	3.122***	0.000
	시야차단 지수 (VOI)	0.000	1.000	-0.426	0.412	2.794*	0.073	-1.320	0.151
	색상복합 지수 (CEI)	-1.982**	0.047	0.296	0.876	0.709	0.796	-4.208	0.463
Constant		1.481	0.584	1.749	0.347	-1.390	0.596	6.019	0.352
Obs.		284		1,109		69		252	
Log likelihood		-710.87		-3926.9		-146.1		-744.35	
Pseudo R <sup>2</sup>		0.459		0.272		0.771		0.495	

\*p<0.1, \*\*p<0.05, \*\*\*p<0.01

첫째, 녹지 지수(GVI)의 양면적 효과는 기존 문헌과 비교해 그 맥락적 함의가 두드러진다. 다수의 선행연구(Nitoslawski et al., 2016; Song et al., 2018; Van Treese et al., 2018)는 도로변 녹지가 운전자 시각적 피로를 완화하고, 주의력 환기를 통해 사고 위험을 감소시키는 효과를 강조하였다. 그러나 일부 연구(Cai et al., 2022; Choi et al., 2020)는 녹지가 오히려 시야를 차단하여 사고 위험을 증가시킬 수 있음을 보고한 바 있다. 이러한 상반된 결과는, 녹지의 배치, 도로 폭, 주변 환경 등 공간적·물리적 맥락에 따라 그 효과가 달라질 수 있음을 시사한다. 본 연구 역시 상업 및 생활 복합구역(Cluster 2)에서는 녹지가 사고 위험을 완화하는 효과가 나타난 반면, 저밀 주거구역(Cluster 4)에서는 녹지가 오히려 사고 위험을 증가시키는 양(+)의 효과를 보였다. 이는 좁은 도로, 무질서한 녹지 배치 등에서 녹지가 시야를 가려 운전자 인지 부하와 사고 위험을 증가시킬 수 있음을 의미하며, 녹지의 효과가 특정 구역에서만 부정적으로 작동하는 공간적 조건을 뚜렷하게 보여준다. 아울러, 녹지 지수는 XGBoost-SHAP 분석에서 높은 변수 중요도를 보였으나, 고밀 주거·교육 복합구역에서 통계적으로 유의하지 않은 결과가 도출되었다. 이는 해당 구역에서 교차로 밀도, 도로 혼잡 등 물리적 요인과의 상호작용 속에서 녹지의 순효과가 상쇄될 수 있음을 보여준다. 따라서, 변수 중요도가 높더라도 통계적 유의성은 공간 맥락에 따라 달라질 수 있으며, 단일 변수의 독립 효과 해석에 유의가 필요하다.

둘째, 색상 복잡 지수(CEI)의 사고 예방 효과는 기능구역의 특수성과 밀접하게 연관된다. 본 연구에서 CEI는 고밀 주거·교육 구역(Cluster 1)에서만 유의한 음(-)의 효과를 보였다. 이는 해당 구역이 대단지 아파트, 교육시설, 공원 등 주거·교육 인프라가 밀집되어 비교적 획일적이고 단조로운 가로경관이 형성되는 곳임을 의미한다. 이처럼 시각적 자극이 적은 환경에서는, 색상 다양성과 대비가 운전자의 각성도와 주의력을 효과적으로 유지시키는 주요한 외부 자극으로 작동할 수 있다(박준상, 2023; Zhang et al., 2023). 반면, 상업·공공서비스나 교통·여가 구역과 같이 본질적으로 다양한 색채·시각정보가 혼재된 환경에서는 색상 복잡도가 추가적인 차별적 인지 부하로 작용하지 않으므로, 효과가 통계적으로 뚜렷하게 나타나지 않은 것으로 해석할 수 있다. 즉, CEI의 효과는 “시각적 단조로움”이 지배적인 구역에서 더욱 두드러지며, 기능구역의 물리적·사회적 맥락에 따라 차별적으로 작동한다.

셋째, 시야 차단지수(VOI)는 교통 및 여가 복합구역(Cluster 3)에서만 유의한 양(+)의 효과를 보였다. 이는 Choi et al.(2020)의 연구에서 지적한 “도로변의 물리적 구조물, 시설물 등 다양한 시야 방해 요소가 운전자의 전방 시야를 제한하여 사고 위험을 증가시킬 수 있다”라는 결과와 일치한다. 본 연구에서 해당 구역은 대형 교통 인프라와 여가시설이 복합적으로 분포하며, 도로 주변에 표지판, 펜스, 방음벽, 고가도로, 전봇대 등 운전자의 시야를

부분적으로 가리는 구조물이 밀집되어 있는 특징을 갖는다. 더불어, 광역 교통도로와 중차량(버스, 트럭 등)의 빈번한 통행이 이루어지는 공간 특성상, 이러한 시야 방해 요소가 더욱 다수 존재할 가능성이 높다. 결과적으로 VOI의 증가는 운전자의 시야 확보를 저해하여, 복잡한 교통·여가 복합 환경에서 교통사고 위험을 실질적으로 증가시키는 핵심 요인으로 작동함을 확인하였다. 이와 같은 결과는 대규모 교통 인프라와 다양한 이용 행태가 중첩되는 도시 공간에서 시야 차단 요소의 효과적 관리가 교통안전 정책 수립 시 반드시 고려되어야 함을 시사한다.

넷째, 각 도시 기능구역 간 SHAP 분석에서 중요도가 높게 나타난 가시적 경관 요소들이 통계적으로 유의하게 작동한 경우가 다수 관찰되었으나, 일부 구역에서는 변수 중요도는 높았으나 통계적으로는 유의하지 않은 사례도 확인되었다. 이러한 불일치는 도시 환경 내 다중 영향요인 및 상호작용의 결과일 수 있으며, 변수의 독립효과가 공간적 맥락에 따라 달라질 수 있음을 시사한다. 이는 SHAP 등 기계학습 기반 변수 중요도 해석과, 전통적 회귀 기반 통계적 검증을 함께 활용할 필요성을 강조하며, 도시공간의 복합성과 이질성을 고려한 해석적 접근이 요구됨을 보여준다.

이상의 논의는, 가시적 경관 요소의 효과가 기능구역의 물리적 구조와 이용 특성에 따라 이질적으로 작동함을 밝혔으며, 선행연구에서 제기된 효과의 불일치와 양면성이 실제 도시공간 내에서 공간적 맥락에 따라 구체적으로 구현됨을 보여준다. 또한, 변수 중요도와 통계적 유의성의 불일치 현상 역시 도시환경의 다중적 영향요인과 상호작용 효과를 고려한 해석의 필요성을 시사한다.

## 2. 연구의 종합 및 정책적 시사점

본 연구는 도시 내 기능구조와 시각환경 요소 간의 상호작용이 교통사고 발생에 미치는 영향을 실증적으로 분석하고자 하였다. 이를 위해 다음의 두 가지 연구질문을 설정하고, 그에 대한 실증 분석을 수행하였다.

첫째, 실제 도시 활동 특성을 반영한 기능구역은 어떠한 유형으로 구분되는가(RQ1)에 대해, 본 연구는 POI 데이터와 임베딩 기법을 활용하여 도시 내 주요 활동 중심지를 정량적으로 계량화하고, K-means 클러스터링 기법을 통해 4개의 도시 기능구역을 도출하였다. 도출된 기능구역은 각각 ① 고밀 주거·교육 중심 구역, ② 생활 복합 구역, ③ 교통·여가 복합 구역, ④ 저밀 주거·생활 구역으로 명명되었으며, 위성 이미지와 주요 시설 분포와의 비교를 통해 해당 분류가 실제 도시 공간의 기능적 특성과 일정 수준의 정합성을 갖는 것을 확인하였다.

둘째, 도시 기능구역별로 가시적 경관 요소는 교통사고 발생에 어떠한 차별적 영향을 미치는가(RQ2)에 대해서는, SHAP와 음이항 회귀모형을 병행 적용하여 시각환경 요소의 중요성과 통계적 유의성을 검토하였다. 분석 결과, 기능구역별로 가시적 경관

요소의 영향력이 상이하게 나타났으며, 특히 녹지 지수(GVI)는 기능구역에 따라 부(-), 정(+), 무(0)의 다양한 효과를 보여주었다. 이는 동일한 시각환경 요소라도 도시 공간 구조 및 기능과 결합하는 방식에 따라 사고에 미치는 영향이 달라질 수 있음을 실증적으로 입증한 결과이다.

따라서 본 연구는 도시경관 설계 및 교통안전 정책 수립에 있어 다음과 같은 기능구역별 맞춤형 접근의 필요성을 제안한다.

첫째, 고밀 주거·교육 중심 구역에서는 단순로운 시각 환경이 운전자의 주의력 저하로 이어질 수 있다. 이에 따라 색채, 재료, 형태 등 시각적 다양성을 확보할 수 있는 경관 디자인이 중요하며, 아파트 단지나 교육시설 주변 도로에 적용 가능한 시각환경 개선 가이드라인의 마련이 요구된다.

둘째, 생활 복합 구역에서는 도심 내 녹지가 시각적 완충지대로 작용하여 사고 위험을 완화하는 데 기여할 수 있다. 따라서 가로 녹지의 적절한 확보와 체계적인 유지관리 기준 강화가 필요하며, 특히 교차로 주변과 보행 통행량이 많은 구간에서는 녹지의 배치, 면적, 시야 확보 여부를 종합적으로 고려한 설계가 중요하다.

셋째, 교통·여가 복합 구역은 여가시설, 공공광장, 교차로 등이 혼재된 복합공간으로, 물리적 구조물(예: 조경 시설, 안내 구조물 등)이 시야를 차단해 사고 위험을 높일 수 있다. 따라서 해당 구역에서는 교차로 및 주요 진입부의 시각적 개방성 확보와 시설물 배치의 정비 및 조정이 핵심적인 안전 전략으로 요구된다.

마지막으로, 저밀 주거·생활 구역은 생활도로 중심의 주거환경으로, 무계획적인 녹지 배치가 운전자의 시야 확보를 방해하고 사고 위험을 증대시킬 수 있다. 이에 따라 식재 간격, 수종 선정, 식재 높이 등 운전자의 시야를 고려한 세부적인 식재계획 수립이 필요하다.

시각환경 요소는 그 자체로 절대적인 효과를 가지기보다 기능구역의 공간적 맥락과 상호작용하며 교통사고 발생에 영향을 미친다. 따라서 기능구역별 특성을 반영한 정밀한 경관 설계 및 맞춤형 교통안전 전략이 향후 도시계획과 정책 수립의 핵심 과제로 부상해야 할 것이다.

### 3. 연구의 한계 및 향후 연구

본 연구는 기능구역별 시각환경 요소가 교통사고에 미치는 차별적 영향을 실증적으로 분석하였으나, 다음과 같은 한계를 가진다.

첫째, 기능구역의 공간 특성 검토가 위성영상 기반 대표 사례 분석에 한정되어 정성적 수준에 머물렀다. 향후에는 건축밀도, 통행량 등 다양한 정량적 공간지표를 활용한 다중 사례 분석을 통해 클러스터 유형의 정합성을 체계적으로 검증할 필요가 있다.

둘째, 카카오맵 POI 15개 분류체계에 근거한 일괄 클러스터링으로 도시 내 세부 기능의 미묘한 차이나 공간적 변이를 충분히 포착하지 못했을 가능성이 있다. 보다 세밀한 기능 분류와 다단계

클러스터링 기법의 도입이 요구된다.

셋째, 교통사고 발생과 관련된 노출변수를 제어하기 위해 교통 혼잡비용을 대리변수로 활용하였다. 이는 실제 통행량을 대표하는 일평균교통량과 달리 혼잡으로 인한 시간 손실 등 다양한 요소가 혼잡된 지표로, 차량 통행이 많은 지역과 단순히 혼잡도가 높은 지역을 명확히 구분하지 못하는 한계가 있다. 특히 국내에서는 일평균교통량 등 세밀한 교통량 자료가 충분히 구축·공개되어 있지 않아, 향후 관련 자료의 구축 및 활용을 통해 통계변수의 정밀도를 제고할 필요가 있다.

마지막으로, 본 연구는 시각환경 요소와 교통사고 간의 통계적 연관성에 초점을 맞추었으나, 실제 사고에 영향을 미치는 인지적·행태적 메커니즘까지는 설명하지 못하였다. 향후 연구에서는 시뮬레이션, 주행 실험, 시선추적 등 다양한 정성·정량 연구방법을 병행하여 시각환경 요소의 작동 경로와 인지적 영향 과정을 심층적으로 규명할 필요가 있다. 이상의 결과는 도시기능구역별 맞춤형 시각환경 관리와 교통안전정책의 중요성을 뒷받침하며, 향후 도시계획·교통안전 분야의 실천적·학술적 논의에 기여할 수 있을 것으로 기대된다.

### 인용문헌 References

- 김선재·이수기, 2021. "POI 빅데이터를 활용한 도시활동 중심지 도출과 중심지 기능 분석: 서울 대도시권을 중심으로", 「국토계획」, 56(6): 36-52.  
Kim, S. and Lee, S., 2023. "Identifying Urban Activity Centers and Their Functions Using POI Big Data: The Case of Seoul Metropolitan Area", *Journal of Korea Planning Association*, 56(6): 36-52.
- 박근덕·기동환·이수기, 2021. "서울시 가로환경 요소의 시각적 특성이 보행만족도에 미치는 영향 분석: 구글 가로이미지와 딥러닝 의미론적 분할 기법을 활용하여", 「도시설계」, 22(3): 55-72.  
Park, K., Ki, D., and Lee, S., 2021. "Analysis of Visual Characteristics of Urban Street Elements on Walking Satisfaction in Seoul, Korea: Application of Google Street View and Deep Learning Technique of Semantic Segmentation", *Journal of the Urban Design Institute of Korea*, 22(3): 55-72.
- 박은총·이수기, 2023. "고령 보행자 교통사고 영향요인의 상호작용 효과 분석: 보행환경개선사업과 보행자안전시설을 중심으로", 「국토계획」, 58(5): 108-124.  
Park, E. and Lee, S., 2023. "Interaction Effect Analysis of Determining Factors for Older Pedestrian Crash: Focused on Walking Environment Improvement Projects and Pedestrian Safety Facilities", *Journal of Korea Planning Association*, 58(5): 108-124.
- 박준상, 2023. "서울시 가로경관의 가시적 특성이 자동차 교통사고에 미치는 영향 분석", 한양대학교 석사학위 논문.

- Park, J., 2023. "Analysis of the Visual Characteristics of Street Scenery Affecting Car Traffic Accidents in Seoul, Korea: Using Street View Panorama Image and Computer Vision Techniques", Master's Thesis, Hanyang University.
5. 서지민·이수기, 2016. "서울시 보행자 교통사고에 영향을 미치는 물리적 환경요인에 관한 연구: 2014 TAAS 자료를 중심으로", 「국토계획」, 51(3): 197-216.  
Seo, J. and Lee, S., 2016. "A Study on the Physical Environmental Factors Influencing Pedestrian Traffic Accidents in Seoul, Korea : Focused on the 2014 TAAS Data", *Journal of Korea Planning Association*, 51(3): 197-216.
  6. 성우석·이수기, 2025. "도시환경 특성이 범죄발생에 미치는 영향 분석: 사회적 약자 그룹의 조절효과를 중심으로", 「국토계획」, 60(1): 5-24.  
Sung, W. and Lee, S., 2025. "Analyzing the Impacts of Urban Environmental Characteristics on Crime Occurrence: Focusing on the Moderating Effects of Socially Disadvantaged Groups", *Journal of Korea Planning Association*, 60(1): 5-24.
  7. 이영민, 2022. 「POI(관심지점) 데이터를 활용한 활동 중심지 추출 및 분석 방법: 세종시를 사례로」, 국토연구원.  
Lee, Y.M., 2022. *Methods for Extracting and Analyzing Activity Centers using POI Data: The Case of Sejong City*, Korea Research Institute for Human Settlements.
  8. 이지원·김지영, 2023. "POI 임베딩 기법에 기반한 도시기능지역 추출방법", 「대한공간정보학회지」, 31(4): 53-63.  
Lee, J.W. and Kim, J.Y., 2023. "Urban Functional Area Extraction Method Based on POI Embedding Technique", *Journal of Korean Society for Geospatial Information Science*, 31(4): 53-63.
  9. 이지원·김태형, 2019. "서울시 고령 운전자 교통사고 특성 분석", 「국토연구」, 102(1): 19-34.  
Lee, J.W. and Gim, T.H., 2019. "Examining the Characteristics of Traffic Accidents Involving Elderly Drivers in Seoul, South Korea", *The Korea Spatial Planning Review*, 102(1): 19-34.
  10. 황선근·이수기, 2021. "서울시 자전거 교통사고의 물리적 환경요인 분석", 「국토계획」, 56(5): 83-96.  
Hwang, S. and Lee, S., 2021. "Analysis of Physical Environment Factors of Bicycle Accidents in Seoul, Korea", *Journal of Korea Planning Association*, 56(5): 83-96.
  11. Aboulola, O.I., 2024. "Improving Traffic Accident Severity Prediction Using MobileNet Transfer Learning Model and SHAP XAI Technique", *PLoS One*, 19(4): e0300640.
  12. Cai, Q., Abdel-Aty, M., Zheng, O., and Wu, Y., 2022. "Applying Machine Learning and Google Street View to Explore Effects of Drivers' Visual Environment on Traffic Safety", *Transportation Research Part C: Emerging Technologies*, 135: 103541.
  13. Chang, I., Park, H., Hong, E., Lee, J., and Kwon, N., 2022. "Predicting Effects of Built Environment on Fatal Pedestrian Accidents at Location-Specific Level: Application of XGBoost and SHAP", *Accident Analysis & Prevention*, 166: 106545.
  14. Chi, J., Jiao, L., Dong, T., Gu, Y., and Ma, Y., 2016. "Quantitative Identification and Visualization of Urban Functional Area Based on POI Data", *Journal of Geometry*, 41(6): 68-73.
  15. Chiang, Y.C., Ke, R.A., Li, D., and Weng, P.Y., 2022. "Greening and Safety: The Influence of Road Greenness on Driver's Attention and Emergency Reaction Time", *Environment and Behavior*, 54(9-10): 1195-1226.
  16. Choi, K., Byun, G., Kim, A., and Kim, Y., 2020. "Drivers' Visual Perception Quantification Using 3D Mobile Sensor Data for Road Safety", *Sensors*, 20(10): 2763.
  17. Dong, N., Huang, H., and Zheng, L., 2015. "Support Vector Machine in Crash Prediction at the Level of Traffic Analysis Zones: Assessing the Spatial Proximity Effects", *Accident Analysis & Prevention*, 82: 192-198.
  18. Dukic, T., Ahlstrom, C., Patten, C., Kettwich, C., and Kircher, K., 2013. "Effects of Electronic Billboards on Driver Distraction", *Traffic Injury Prevention*, 14(5): 469-476.
  19. Fu, R., Guo, Y., Yuan, W., Feng, H., and Ma, Y., 2011. "The Correlation Between Gradients of Descending Roads and Accident Rates", *Safety Science*, 49(3): 416-423.
  20. Han, Y., 2025. "Investigating Streetscape Environmental Characteristics Associated with Road Traffic Crashes Using Street View Imagery and Computer Vision", *Accident Analysis and Prevention*, 210: 107851.
  21. Harvey, C. and Aultman-Hall, L., 2015. "Urban Streetscape Design and Crash Severity", *Transportation Research Record: Journal of the Transportation Research Board*, 2500(1): 1-8.
  22. Kang, Y., Wang, Y., Xia, Z., Chi, J., Jiao, L., and Wei, Z., 2018. "Identification and Classification of Wuhan Urban Districts Based on POI", *Journal of Geometry*, 43(1): 81-85.
  23. Kwon, J.H. and Cho, G.H., 2020. "An Examination of the Intersection Environment Associated with Perceived Crash Risk among School-Aged Children: Using Street-Level Imagery and Computer Vision", *Accident Analysis & Prevention*, 146: 105716.
  24. Lin, A., Wu, H., Luo, W., Fan, K., and Liu, H., 2024. "How Does Urban Heat Island Differ Across Urban Functional Zones Insights from 2D3D Urban Morphology Using Geospatial Big Data", *Urban Climate*, 53: 101787.
  25. Liu, B., Deng, Y., Li, M., Yang, J., and Liu, T., 2021. "Classification Schemes and Identification Methods for Urban Functional Zone: A Review of Recent Papers", *Applied Sciences*, 11(21): 9968.
  26. Liu, W., Zhang, X., Feng, Q., Yu, T., and Engel, B.A., 2023. "Analyzing The Impacts of Topographic Factors and Land Cover Characteristics on Waterlogging Events in Urban Functional Zones", *Science of The Total Environment*, 904: 166669.
  27. Liu, X., Wu, T., Jiang, Q., Ao, X., Zhu, L., and Qiao, R., 2025. "The Nonlinear Climatological Impacts of Urban Morphology on Extreme Heats in Urban Functional Zones: An Interpretable Ensemble Learning-Based Approach", *Building and Environment*, 273: 112728.
  28. Nassiri, H., Mohammadpour, S.I., and Dahaghin, M., 2023. "Forecasting Time Trends of Fatal Motor Vehicle Crashes in Iran Using an Ensemble Learning Algorithm", *Traffic Injury*

- Prevention*, 24(1): 44-49.
29. Nitoslowski, S.A., Duinker, P.N., and Bush, P.G., 2016. "A Review of Drivers of Tree Diversity in Suburban Areas: Research Needs for North American Cities", *Environmental Reviews*, 24(4): 471-483.
  30. Parsa, A.B., Movahedi, A., Taghipour, H., Derrible, S., and Mohammadian, A.K., 2020. "Toward Safer Highways, Application of XGBoost and SHAP for Real-Time Accident Detection and Feature Analysis", *Accident Analysis & Prevention*, 136: 105405.
  31. Regan, M.A., Hallett, C., and Gordon, C.P., 2011. "Driver Distraction and Driver Inattention: Definition, Relationship and Taxonomy", *Accident Analysis & Prevention*, 43(5): 1771-1781.
  32. Song, X.P., Tan, P.Y., Edwards, P., and Richards, D., 2018. "The Economic Benefits And Costs of Trees in Urban Forest Stewardship: A Systematic Review", *Urban Forestry & Urban Greening*, 29: 162-170.
  33. Stiles, J., Li, Y., and Miller, H.J., 2022. "How Does Street Space Influence Crash Frequency? An Analysis Using Segmented Street View Imagery", *Environment and Planning B: Urban Analytics and City Science*, 49(9): 2467-2483.
  34. Sun, Z., Wang, J., Chen, Y., and Lu, H., 2018. "Influence Factors on Injury Severity of Traffic Accidents and Differences in Urban Functional Zones: The Empirical Analysis of Beijing", *International Journal of Environmental Research and Public Health*, 15(12): 2722.
  35. Van Treese, J., Koeser, A.K., Fitzpatrick, G.E., Olexa, M.T., and Allen, E.J., 2018. "Drivers' Risk Perception of Roadside Trees", *Arboricultural Journal*, 40(3): 153-161.
  36. Wang, H., Li, P., Liu, Y., and Shao, J., 2021. "Towards Real-Time Demand-Aware Sequential POI Recommendation", *Information Sciences*, 547: 482-497.
  37. Yan, B., Janowicz, K., Mai, G., and Gao, S., 2017. "From ITDL to Place2Vec: Reasoning About Place Type Similarity and Relatedness by Learning Embeddings from Augmented Spatial Contexts", *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 35: 1-10.
  38. Yang, C., Chen, M., and Yuan, Q., 2021. "The Application of XGBoost and SHAP to Examining the Factors in Freight Truck-Related Crashes: An Exploratory Analysis", *Accident Analysis & Prevention*, 158: 106153.
  39. Ye, X., Wang, K., Zou, Y., and Lord, D., 2018. "A Semi-Non-parametric Poisson Regression Model for Analyzing Motor Vehicle Crash Data", *PLoS One*, 13(5): e0197338.
  40. Yuan, Y., Chen, X., and Wang, J., 2020. "Object-Contextual Representations for Semantic Segmentation", *Proceedings of the 16th European Conference on Computer Vision*, 173-190, Scottish Event Campus, Glasgow, U.K., August 23-28.
  41. Yue, H., 2024. "Investigating the Influence of Streetscape Environmental Characteristics on Pedestrian Crashes at Intersections Using Street View Images and Explainable Machine Learning", *Accident Analysis & Prevention*, 205: 107693.
  42. Zhai, W., Bai, X., Shi, Y., Han, Y., Peng, Z. R., and Gu, C., 2019. "Beyond Word2vec: An Approach for Urban Functional Region Extraction and Identification by Combining Place2vec and POIs", *Computers, Environment and Urban Systems*, 74: 1-12.
  43. Zhu, X. and Zhou, C., 2009. "POI Inquiries and Data Update Based on LBS", *Proceedings of the 2009 International Symposium on Information Engineering and Electronic Commerce*, 730-734.
  44. Zhang, H., Hou, N., Ding, N., and Jiao, N., 2023. "Using Multicolor Perceptual Markings as A Rear-End Crash Risk Mitigator: A field Investigation", *Accident Analysis & Prevention*, 179: 106881.

Date Received 2025-06-24  
 Date Reviewed 2025-07-25  
 Date Accepted 2025-07-25  
 Date Revised 2025-08-28  
 Final Received 2025-08-28